

# Lecture Notes<sup>1</sup>

## Microeconomic Theory

Guoqiang TIAN  
Department of Economics  
Texas A&M University  
College Station, Texas 77843  
(gtian@tamu.edu)

August, 2002/Revised: February 2013

---

<sup>1</sup>This lecture notes are only for the purpose of my teaching and convenience of my students in class, but not for any other purpose.

# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Preliminaries on Modern Economics and Mathematics</b>   | <b>1</b>  |
| 1.1      | Nature of Modern Economics . . . . .   | 1         |
| 1.1.1    | Modern Economics and Economic Theory . . . . .   | 1         |
| 1.1.2    | Key Assumptions and Desired Properties Commonly Used Economics   | 3         |
| 1.1.3    | The Basic Analytical Framework of Modern Economics . . . . .   | 4         |
| 1.1.4    | Methodologies for Studying Modern Economics . . . . .  | 6         |
| 1.1.5    | Roles, Generality, and Limitation of Economic Theory . . . . .   | 8         |
| 1.1.6    | Roles of Mathematics in Modern Economics . . . . .   | 10        |
| 1.1.7    | Conversion between Economic and Mathematical Languages . . . .   | 11        |
| 1.1.8    | Distinguish between Necessary and Sufficient Conditions as well<br>Normative and Positive Statements . . . . . | 11        |
| 1.2      | Language and Methods of Mathematics . . . . .  | 12        |
| 1.2.1    | Functions . . . . .  | 12        |
| 1.2.2    | Separating Hyperplane Theorem . . . . .  | 14        |
| 1.2.3    | Concave and Convex Functions . . . . .   | 14        |
| 1.2.4    | Optimization . . . . .   | 16        |
| 1.2.5    | The Envelope Theorem . . . . .   | 19        |
| 1.2.6    | Point-to-Set Mappings . . . . .  | 21        |
| 1.2.7    | Continuity of a Maximum . . . . .  | 25        |
| 1.2.8    | Fixed Point Theorems . . . . .   | 26        |
| <b>I</b> | <b>Individual Decision Making</b>  | <b>30</b> |
| <b>2</b> | <b>Consumer Theory</b>   | <b>32</b> |

|        |   |    |
|--------|---|----|
| 2.1    | Introduction . . . . .  | 32 |
| 2.2    | Consumption Set and Budget Constraint . . . . .                                 | 33 |
| 2.2.1  | Consumption Set . . . . .   | 33 |
| 2.2.2  | Budget Constraint . . . . .   | 33 |
| 2.3    | Preferences and Utility . . . . .   | 35 |
| 2.3.1  | Preferences . . . . .   | 35 |
| 2.3.2  | The Utility Function . . . . .  | 40 |
| 2.4    | Utility Maximization and Optimal Choice . . . . .                               | 45 |
| 2.4.1  | Consumer Behavior: Utility Maximization . . . . .                               | 45 |
| 2.4.2  | Consumer's Optimal Choice . . . . .   | 45 |
| 2.4.3  | Consumer's First Order-Conditions . . . . .                                     | 46 |
| 2.4.4  | Sufficiency of Consumer's First-Order Conditions . . . . .                      | 49 |
| 2.5    | Indirect Utility, and Expenditure, and Money Metric Utility Functions . . . . . | 52 |
| 2.5.1  | The Indirect Utility Function . . . . .   | 52 |
| 2.5.2  | The Expenditure Function and Hicksian Demand . . . . .                          | 54 |
| 2.5.3  | The Money Metric Utility Functions . . . . .                                    | 57 |
| 2.5.4  | Some Important Identities . . . . .   | 60 |
| 2.6    | Duality Between Direct and Indirect Utility . . . . .                           | 64 |
| 2.7    | Properties of Consumer Demand . . . . .   | 66 |
| 2.7.1  | Income Changes and Consumption Choice . . . . .                                 | 66 |
| 2.7.2  | Price Changes and Consumption Choice . . . . .                                  | 66 |
| 2.7.3  | Income-Substitution Effect: The Slutsky Equation . . . . .                      | 67 |
| 2.7.4  | Continuity and Differentiability of Demand Functions . . . . .                  | 70 |
| 2.7.5  | Inverse Demand Functions . . . . .  | 72 |
| 2.8    | The Integrability Problem . . . . .   | 73 |
| 2.9    | Revealed Preference . . . . .   | 76 |
| 2.9.1  | Axioms of Revealed Preferences . . . . .  | 76 |
| 2.9.2  | Characterization of Revealed Preference Maximization . . . . .                  | 78 |
| 2.10   | Recoverability . . . . .  | 80 |
| 2.11   | Topics in Demand Behavior . . . . .   | 82 |
| 2.11.1 | Endowments in the Budget Constraint . . . . .                                   | 82 |

|          |   |           |
|----------|---|-----------|
| 2.11.2   | Income-Leisure Choice Model . . . . .                                 | 83        |
| 2.11.3   | Homothetic Utility Functions . . . . .                                | 84        |
| 2.11.4   | Aggregating Across Goods . . . . .                                    | 84        |
| 2.11.5   | Aggregating Across Consumers . . . . .                                | 90        |
| <b>3</b> | <b>Production Theory</b>  | <b>95</b> |
| 3.1      | Introduction . . . . .  | 95        |
| 3.2      | Production Technology . . . . .                                       | 96        |
| 3.2.1    | Measurement of Inputs and Outputs . . . . .                           | 96        |
| 3.2.2    | Specification of Technology . . . . .                                 | 96        |
| 3.2.3    | Common Properties of Production Sets . . . . .                        | 100       |
| 3.2.4    | Returns to Scale . . . . .  | 102       |
| 3.2.5    | The Marginal Rate of Technical Substitution . . . . .                 | 103       |
| 3.2.6    | The Elasticity of Substitution . . . . .                              | 104       |
| 3.3      | Profit Maximization . . . . .   | 105       |
| 3.3.1    | Producer Behavior . . . . .   | 105       |
| 3.3.2    | Producer's Optimal Choice . . . . .                                   | 107       |
| 3.3.3    | Producer's First-Order Conditions . . . . .                           | 108       |
| 3.3.4    | Sufficiency of Producer's First-Order Condition . . . . .             | 109       |
| 3.3.5    | Properties of Net Supply Functions . . . . .                          | 111       |
| 3.3.6    | Weak Axiom of Profit Maximization . . . . .                           | 112       |
| 3.3.7    | Recoverability . . . . .  | 113       |
| 3.4      | Profit Function . . . . .   | 115       |
| 3.4.1    | Properties of the Profit Function . . . . .                           | 116       |
| 3.4.2    | Deriving Net Supply Functions from Profit Function . . . . .          | 117       |
| 3.5      | Cost Minimization . . . . .   | 118       |
| 3.5.1    | First-Order Conditions of Cost Minimization . . . . .                 | 119       |
| 3.5.2    | Sufficiency of First-Order Conditions for Cost Minimization . . . . . | 120       |
| 3.6      | Cost Functions . . . . .  | 124       |
| 3.6.1    | Properties of Cost Functions . . . . .                                | 124       |
| 3.6.2    | Properties of Conditional Input Demand . . . . .                      | 125       |
| 3.6.3    | Average and Marginal Costs . . . . .                                  | 126       |

|           |  |            |
|-----------|--|------------|
| 3.6.4     | The Geometry of Costs . . . . .                            | 128        |
| 3.6.5     | Long-Run and Short-Run Cost Curves . . . . .               | 129        |
| 3.7       | Duality in Production . . . . .                            | 130        |
| 3.7.1     | Recovering a Production Set from a Cost Function . . . . . | 131        |
| 3.7.2     | Characterization of Cost Functions . . . . .               | 135        |
| 3.7.3     | The Integrability for Cost Functions . . . . .             | 136        |
| <b>4</b>  | <b>Choice Under Uncertainty</b>                            | <b>139</b> |
| 4.1       | Introduction . . . . .                                     | 139        |
| 4.2       | Expected Utility Theory . . . . .                          | 140        |
| 4.2.1     | Lotteries . . . . .  | 140        |
| 4.2.2     | Expected Utility . . . . .                                 | 141        |
| 4.2.3     | Uniqueness of the Expected Utility Function . . . . .      | 144        |
| 4.2.4     | Other Notations for Expected Utility . . . . .             | 145        |
| 4.3       | Risk aversion . . . . .                                    | 145        |
| 4.3.1     | Absolute Risk Aversion . . . . .                           | 145        |
| 4.3.2     | Global Risk Aversion . . . . .                             | 148        |
| 4.3.3     | Relative Risk Aversion . . . . .                           | 152        |
| 4.4       | State Dependent Utility . . . . .                          | 153        |
| 4.5       | Subjective Probability Theory . . . . .                    | 154        |
| <b>II</b> | <b>Strategic Behavior and Markets</b>                      | <b>159</b> |
| <b>5</b>  | <b>Game Theory</b>   | <b>161</b> |
| 5.1       | Introduction . . . . .                                     | 161        |
| 5.2       | Description of a game . . . . .                            | 162        |
| 5.2.1     | Strategic Form . . . . .                                   | 162        |
| 5.3       | Solution Concepts . . . . .                                | 166        |
| 5.3.1     | Mixed Strategies and Pure Strategies . . . . .             | 166        |
| 5.3.2     | Nash equilibrium . . . . .                                 | 167        |
| 5.3.3     | Dominant strategies . . . . .                              | 171        |
| 5.4       | Repeated games . . . . .                                   | 171        |

|            |   |            |
|------------|---|------------|
| 5.5        | Refinements of Nash equilibrium . . . . .                                   | 174        |
| 5.5.1      | Elimination of dominated strategies . . . . .                               | 174        |
| 5.5.2      | Sequential Games and Subgame Perfect Equilibrium . . . . .                  | 175        |
| 5.5.3      | Repeated games and subgame perfection . . . . .                             | 180        |
| 5.6        | Games with incomplete information . . . . .                                 | 181        |
| 5.6.1      | Bayes-Nash Equilibrium . . . . .  | 181        |
| 5.6.2      | Discussion of Bayesian-Nash equilibrium . . . . .                           | 184        |
| <b>6</b>   | <b>Theory of the Market</b>   | <b>186</b> |
| 6.1        | Introduction . . . . .  | 186        |
| 6.2        | The Role of Prices . . . . .  | 187        |
| 6.3        | Perfect Competition . . . . .   | 187        |
| 6.3.1      | Assumptions on Competitive Market . . . . .                                 | 188        |
| 6.3.2      | The Competitive Firm . . . . .  | 188        |
| 6.3.3      | The Competitive Firm's Short-Run Supply Function . . . . .                  | 188        |
| 6.3.4      | Partial Market Equilibrium . . . . .  | 190        |
| 6.3.5      | Competitive in the Long Run . . . . .                                       | 192        |
| 6.4        | Pure Monopoly . . . . .   | 193        |
| 6.4.1      | Profit Maximization Problem of Monopolist . . . . .                         | 193        |
| 6.4.2      | Inefficiency of Monopoly . . . . .  | 195        |
| 6.4.3      | Monopoly in the Long Run . . . . .  | 196        |
| 6.5        | Monopolistic Competition . . . . .  | 197        |
| 6.6        | Oligopoly . . . . .   | 199        |
| 6.6.1      | Cournot Oligopoly . . . . .   | 200        |
| 6.6.2      | Stackelberg Model . . . . .   | 201        |
| 6.6.3      | Bertrand Model . . . . .  | 202        |
| 6.6.4      | Collusion . . . . .   | 204        |
| 6.7        | Monopsony . . . . .   | 205        |
| <b>III</b> | <b>General Equilibrium Theory and Social Welfare</b>                        | <b>207</b> |
| <b>7</b>   | <b>Positive Theory of Equilibrium: Existence, Uniqueness, and Stability</b> | <b>209</b> |

|          |   |            |
|----------|---|------------|
| 7.1      | Introduction . . . . .  | 209        |
| 7.2      | The Structure of General Equilibrium Model . . . . .                      | 211        |
| 7.2.1    | Economic Environments . . . . .   | 211        |
| 7.2.2    | Institutional Arrangement: Private Market Mechanism . . . . .             | 213        |
| 7.2.3    | Individual Behavior Assumptions: . . . . .                                | 214        |
| 7.2.4    | Competitive Equilibrium . . . . .   | 214        |
| 7.3      | Some Examples of GE Models: Graphical Treatment . . . . .                 | 216        |
| 7.3.1    | Pure Exchange Economies . . . . .   | 216        |
| 7.3.2    | The One-Consumer and One Producer Economy . . . . .                       | 222        |
| 7.4      | The Existence of Competitive Equilibrium . . . . .                        | 225        |
| 7.4.1    | The Existence of CE for Aggregate Excess Demand Functions . . . . .       | 226        |
| 7.4.2    | The Existence of CE for Aggregate Excess Demand Correspondences . . . . . | 245        |
| 7.4.3    | The Existence of CE for General Production Economies . . . . .            | 247        |
| 7.5      | The Uniqueness of Competitive Equilibria . . . . .                        | 247        |
| 7.6      | Stability of Competitive Equilibrium . . . . .                            | 251        |
| 7.7      | Abstract Economy . . . . .  | 256        |
| 7.7.1    | Equilibrium in Abstract Economy . . . . .                                 | 257        |
| 7.7.2    | The Existence of Equilibrium for General Preferences . . . . .            | 259        |
| <b>8</b> | <b>Normative Theory of Equilibrium: Its Welfare Properties</b>            | <b>264</b> |
| 8.1      | Introduction . . . . .  | 264        |
| 8.2      | Pareto Efficiency of Allocation . . . . .                                 | 265        |
| 8.3      | The First Fundamental Theorem of Welfare Economics . . . . .              | 271        |
| 8.4      | Calculations of Pareto Optimum by First-Order Conditions . . . . .        | 273        |
| 8.4.1    | Exchange Economies . . . . .  | 274        |
| 8.4.2    | Production Economies . . . . .  | 275        |
| 8.5      | The Second Fundamental Theorem of Welfare Economics . . . . .             | 276        |
| 8.6      | Non-Convex Production Technologies and Marginal Cost Pricing . . . . .    | 282        |
| 8.7      | Pareto Optimality and Social Welfare Maximization . . . . .               | 285        |
| 8.7.1    | Social Welfare Maximization for Exchange Economies . . . . .              | 286        |
| 8.7.2    | Welfare Maximization in Production Economy . . . . .                      | 287        |
| 8.8      | Political Overtones . . . . .   | 289        |

|           |  |            |
|-----------|--|------------|
| <b>9</b>  | <b>Economic Core, Fair Allocations, and Social Choice Theory</b>                     | <b>292</b> |
| 9.1       | Introduction . . . . .   | 292        |
| 9.2       | The Core of Exchange Economies . . . . .   | 293        |
| 9.3       | Fairness of Allocation . . . . .   | 298        |
| 9.4       | Social Choice Theory . . . . .   | 303        |
| 9.4.1     | Introduction . . . . .   | 303        |
| 9.4.2     | Basic Settings . . . . .   | 304        |
| 9.4.3     | Arrow's Impossibility Theorem . . . . .  | 306        |
| 9.4.4     | Some Positive Result: Restricted Domain . . . . .                                    | 309        |
| 9.4.5     | Gibbard-Satterthwaite Impossibility Theorem . . . . .                                | 311        |
| <b>10</b> | <b>General Equilibrium Under Uncertainty</b>   | <b>318</b> |
| 10.1      | Introduction . . . . .   | 318        |
| 10.2      | A Market Economy with Contingent Commodities . . . . .                               | 319        |
| 10.3      | Arrow-Debreu Equilibrium . . . . .   | 323        |
| 10.4      | Sequential Trade . . . . .   | 325        |
| 10.5      | Incomplete Markets . . . . .   | 331        |
| <b>IV</b> | <b>Externalities and Public Goods</b>  | <b>336</b> |
| <b>11</b> | <b>Externalities</b>   | <b>339</b> |
| 11.1      | Introduction . . . . .   | 339        |
| 11.2      | Consumption Externalities . . . . .  | 340        |
| 11.3      | Production Externality . . . . .   | 347        |
| 11.4      | Solutions to Externalities . . . . .   | 349        |
| 11.4.1    | Pigovian Tax . . . . .   | 350        |
| 11.4.2    | Coase' Voluntary Negotiation and Enforceable Property Rights Ap-<br>proach . . . . . | 351        |
| 11.4.3    | Missing Market . . . . .   | 361        |
| 11.4.4    | The Compensation Mechanism . . . . .   | 362        |
| <b>12</b> | <b>Public Goods</b>  | <b>367</b> |
| 12.1      | Introduction . . . . .   | 367        |



|        |   |     |
|--------|---|-----|
| 12.2   | Notations and Basic Settings . . . . .        | 367 |
| 12.3   | Discrete Public Goods . . . . .               | 369 |
| 12.3.1 | Efficient Provision of Public Goods . . . . . | 369 |
| 12.3.2 | Free-Rider Problem . . . . .                  | 370 |
| 12.3.3 | Voting for a Discrete Public Good . . . . .   | 371 |
| 12.4   | Continuous Public Goods . . . . .             | 372 |
| 12.4.1 | Efficient Provision of Public Goods . . . . . | 372 |
| 12.4.2 | Lindahl Equilibrium . . . . .                 | 374 |
| 12.4.3 | Free-Rider Problem . . . . .                  | 378 |

## **V Information, Incentives, Mechanism Design, and Contract Theory 382**

|           |   |
|-----------|---|
| <b>13</b> | <b>Optimal Mechanism Design: Contracts with One-Agent and Hidden Information 387</b>  |
| 13.1      | Introduction . . . . . 387  |
| 13.2      | Basic Settings of Principal-Agent Model with Adverse Selection . . . . . 389          |
| 13.2.1    | Economic Environment (Technology, Preferences, and Information) 389                   |
| 13.2.2    | Contracting Variables: Outcomes . . . . . 389   |
| 13.2.3    | Timing . . . . . 389  |
| 13.3      | The Complete Information Optimal Contract(Benchmark Case) . . . . . 390               |
| 13.3.1    | First-Best Production Levels . . . . . 390  |
| 13.3.2    | Implementation of the First-Best . . . . . 391  |
| 13.3.3    | A Graphical Representation of the Complete Information Optimal Contract . . . . . 392 |
| 13.4      | Incentive Feasible Contracts . . . . . 393  |
| 13.4.1    | Incentive Compatibility and Participation . . . . . 393                               |
| 13.4.2    | Special Cases . . . . . 394   |
| 13.4.3    | Monotonicity Constraints . . . . . 395  |
| 13.5      | Information Rents . . . . . 395   |
| 13.6      | The Optimization Program of the Principal . . . . . 396                               |
| 13.7      | The Rent Extraction-Efficiency Trade-Off . . . . . 397                                |

|           |  |            |
|-----------|--|------------|
| 13.7.1    | The Optimal Contract Under Asymmetric Information . . . . .          | 397        |
| 13.7.2    | A Graphical Representation of the Second-Best Outcome . . . . .      | 400        |
| 13.7.3    | Shutdown Policy . . . . .  | 400        |
| 13.8      | The Theory of the Firm Under Asymmetric Information . . . . .        | 402        |
| 13.9      | Asymmetric Information and Marginal Cost Pricing . . . . .           | 402        |
| 13.10     | The Revelation Principle . . . . .                                   | 403        |
| 13.11     | A More General Utility Function for the Agent . . . . .              | 405        |
| 13.11.1   | The Optimal Contract . . . . .                                       | 405        |
| 13.11.2   | More than One Good . . . . .   | 407        |
| 13.12     | Ex Ante versus Ex Post Participation Constraints . . . . .           | 408        |
| 13.12.1   | Risk Neutrality . . . . .  | 408        |
| 13.12.2   | Risk Aversion . . . . .  | 410        |
| 13.13     | Commitment . . . . .   | 414        |
| 13.13.1   | Renegotiating a Contract . . . . .                                   | 414        |
| 13.13.2   | Reneging on a Contract . . . . .                                     | 415        |
| 13.14     | Informative Signals to Improve Contracting . . . . .                 | 415        |
| 13.14.1   | Ex Post Verifiable Signal . . . . .                                  | 416        |
| 13.14.2   | Ex Ante Nonverifiable Signal . . . . .                               | 417        |
| 13.15     | Contract Theory at Work . . . . .                                    | 417        |
| 13.15.1   | Regulation . . . . .   | 418        |
| 13.15.2   | Nonlinear Pricing by a Monopoly . . . . .                            | 418        |
| 13.15.3   | Quality and Price Discrimination . . . . .                           | 419        |
| 13.15.4   | Financial Contracts . . . . .  | 420        |
| 13.15.5   | Labor Contracts . . . . .  | 421        |
| 13.16     | The Optimal Contract with a Continuum of Types . . . . .             | 422        |
| 13.17     | Further Extensions . . . . .   | 426        |
| <b>14</b> | <b>Optimal Mechanism Design: Contracts with One-Agent and Hidden</b> |            |
|           | <b>Action</b>  | <b>429</b> |
| 14.1      | Introduction . . . . .   | 429        |
| 14.2      | Basic Settings of Principal-Agent Model with Moral Hazard . . . . .  | 430        |
| 14.2.1    | Effort and Production . . . . .                                      | 430        |

|           |  |            |
|-----------|--|------------|
| 14.2.2    | Incentive Feasible Contracts . . . . .                                   | 431        |
| 14.2.3    | The Complete Information Optimal Contract . . . . .                      | 432        |
| 14.3      | Risk Neutrality and First-Best Implementation . . . . .                  | 433        |
| 14.4      | The Trade-Off Between Limited Liability Rent Extraction and Efficiency . | 435        |
| 14.5      | The Trade-Off Between Insurance and Efficiency . . . . .                 | 437        |
| 14.5.1    | Optimal Transfers . . . . .  | 437        |
| 14.5.2    | The Optimal Second-Best Effort . . . . .                                 | 439        |
| 14.6      | More than Two Levels of Performance . . . . .                            | 440        |
| 14.6.1    | Limited Liability . . . . .  | 440        |
| 14.6.2    | Risk Aversion . . . . .  | 442        |
| 14.7      | Contract Theory at Work . . . . .  | 443        |
| 14.7.1    | Efficiency Wage . . . . .  | 443        |
| 14.7.2    | Sharecropping . . . . .  | 444        |
| 14.7.3    | Wholesale Contracts . . . . .  | 446        |
| 14.7.4    | Financial Contracts . . . . .  | 446        |
| 14.8      | A Continuum of Performances . . . . .                                    | 448        |
| 14.9      | Further Extension . . . . .  | 449        |
| <b>15</b> | <b>General Mechanism Design: Contracts with Multi-Agents</b>             | <b>452</b> |
| 15.1      | Introduction . . . . .   | 452        |
| 15.2      | Basic Settings . . . . .   | 454        |
| 15.2.1    | Economic Environments . . . . .  | 454        |
| 15.2.2    | Social Goal . . . . .  | 455        |
| 15.2.3    | Economic Mechanism . . . . .   | 456        |
| 15.2.4    | Solution Concept of Self-Interested Behavior . . . . .                   | 457        |
| 15.2.5    | Implementation and Incentive Compatibility . . . . .                     | 458        |
| 15.3      | Examples . . . . .   | 460        |
| 15.4      | Dominant Strategy and Truthful Revelation Mechanisms . . . . .           | 461        |
| 15.5      | Gibbard-Satterthwaite Impossibility Theorem . . . . .                    | 464        |
| 15.6      | Hurwicz Impossibility Theorem . . . . .                                  | 465        |
| 15.7      | Vickrey-Clark-Groves Mechanisms . . . . .                                | 468        |
| 15.7.1    | Vickrey-Clark-Groves Mechanisms for Discrete Public Good . . . .         | 468        |

|           |  |            |
|-----------|--|------------|
| 15.7.2    | The Vickrey-Clark-Groves Mechanisms with Continuous Public Goods . . . . .       | 472        |
| 15.7.3    | Balanced VCG Mechanisms . . . . .  | 477        |
| 15.8      | Nash Implementation . . . . .  | 479        |
| 15.8.1    | Nash Equilibrium and General Mechanism Design . . . . .                          | 479        |
| 15.8.2    | Characterization of Nash Implementation . . . . .                                | 481        |
| 15.9      | Better Mechanism Design . . . . .  | 487        |
| 15.9.1    | Groves-Ledyard Mechanism . . . . .   | 487        |
| 15.9.2    | Walker's Mechanism . . . . .   | 489        |
| 15.9.3    | Tian's Mechanism . . . . .   | 491        |
| 15.10     | Incomplete Information and Bayesian-Nash Implementation . . . . .                | 494        |
| 15.10.1   | Bayesian-Nash Implementation Problem . . . . .                                   | 494        |
| 15.10.2   | Ex-Post Efficient Implementation by the Expected Externality Mechanism . . . . . | 497        |
| 15.10.3   | Participation Constraints . . . . .  | 501        |
| 15.10.4   | The Revenue Equivalence Theorem in Auctions . . . . .                            | 506        |
| 15.10.5   | Correlated Types . . . . .   | 510        |
| 15.10.6   | Ex Post Implementation . . . . .   | 513        |
| 15.10.7   | Ex post Implementation with Common Values . . . . .                              | 514        |
| <b>16</b> | <b>Dynamic Mechanism Design</b>  | <b>523</b> |
| 16.1      | Dynamic Contracts with Commitment . . . . .                                      | 523        |
| 16.1.1    | Dynamic P-A Model with Constant Type . . . . .                                   | 524        |
| 16.1.2    | Dynamic P-A Model with Markov Types . . . . .                                    | 525        |
| 16.2      | Dynamic Contracts without Full Commitment . . . . .                              | 527        |
| 16.2.1    | Basic Problem . . . . .  | 527        |
| 16.2.2    | Full Analysis of A Simple Model . . . . .  | 530        |

# Chapter 1

## Preliminaries on Modern Economics and Mathematics

In this chapter, we first set out some basic terminologies and key assumptions imposed in modern economics in general and in the lecture notes in particular. We will discuss the standard analytical framework usually used in modern economics. We will also discuss methodologies for studying modern economics as well as some key points one should pay attention. The methodologies for studying modern economics and key points include: providing studying platforms, establishing reference/benchmark systems, developing analytical tools, noting generality and limitation of an economic theory, the role of mathematics, distinguishing necessary and sufficient conditions for a statement, and conversion between economic and mathematical language. We will then discuss some basic mathematics results that we will use in the lecture notes.

### 1.1 Nature of Modern Economics

#### 1.1.1 Modern Economics and Economic Theory

- What is economics about?

*Economics* is a social science that studies individuals' economic behavior, economic phenomena, as well as how individual agents, such as consumers, firms, and government agencies, make trade-off choices that allocate limited resources among competing uses.

People's desires are unlimited, but resources are limited, therefore individuals must make trade-offs. We need economics to study this fundamental conflict and how these trade-offs are best made.

- Four basic questions must be answered by any economic institution:

- (1) What goods and services should be produced and in what quantity?
- (2) How should the product be produced?
- (3) For whom should it be produced and how should it be distributed?
- (4) Who makes the decision?

The answers depend on the use of economic institutions. There are two basic economic institutions that have been so far used in the real world:

- (1) Market economic institution (the price mechanism): Most decisions on economic activities are made by individuals. This primarily decentralized decision system is the most important economic institution discovered for reaching cooperation amongst individuals and solving the conflicts that occur between them. The market economy has been proven to be only economic institution, so far, that can keep sustainable development and growth within an economy.
- (2) Planned economic institution: Most decisions on economic activities are made by governments, which are mainly centralized decision systems.

- What is Modern Economics?

*Modern economics*, mainly developed in last sixty years, systematically studies individuals' economic behavior and economic phenomena by a scientific studying method – observation → theory → observation – and through the use of various analytical approaches.

- What is Economic Theory?

An *economic theory*, which can be considered an axiomatic approach, consists of a set of assumptions and conditions, an analytical framework, and conclusions (explanations and/or predications) that are derived from the assumptions and the analytical framework.

Like any science, economics is concerned with the explanation of observed phenomena and also makes economic predictions and assessments based on economic theories. Economic theories are developed to explain the observed phenomena in terms of a set of basic assumptions and rules.

- Microeconomic theory

Microeconomic theory aims to model economic activities as the interaction of individual economic agents pursuing their private interests.

### **1.1.2 Key Assumptions and Desired Properties Commonly Used Economics**

Economists usually make all or some of the following key assumptions and conditions when they study economic problems:

- (1) Individuals are (bounded) rational: self-interested behavior assumption;
- (2) Scarcity of resources: individuals confront scarce resources;
- (3) Information about individuals's economic characteristics and actions is incomplete or asymmetric to a decision maker. As such decentralized decision makings is desired.
- (4) Economic freedom: voluntary cooperation and voluntary exchange;
- (5) Incentive compatibility of parties: the system or economic mechanism should solve the problem of interest conflicts among individuals or economic units;
- (6) Well-defined property rights;
- (7) Equity in opportunity;
- (8) Allocative efficiency of resources;

Relaxing any of these assumptions may result in different conclusions.

### **1.1.3 The Basic Analytical Framework of Modern Economics**

The basic analytical framework for an economic theory consists of five aspects or steps: (1) specification of economic environments, (2) imposition of behavioral assumptions, (3) adoption of economic institutional arrangements, (4) determination of equilibria, and (5) evaluation of outcomes resulting from a undertaken institution. The framework is a framework that uses to deal with daily activities and is used to study particular economic issues and questions that economists are interested in. Almost of all theoretical economics papers adopt this framework. As such, to have a solid training in economics and doing research, it is very important to master this basic analytical framework, specifically, these five steps. It can be helpful at least in the following three aspects: (1) Help to understand economic theories and their arguments relatively easily. (2) Help to find research topics. (3) How to write standard scientific economics papers.

Understanding this basic analytical framework can help people classify possible misunderstandings about modern economics, and can also help them use the basic economic principles or develop new economic theories to solve economic problems in various economic environments, with different human behavior and institutional arrangements.

#### **1. Specification of Economic Environments**

The first step for studying an economic issue is to specify the economic environment. The specification on economic environment can be divided into two levels: 1) description of the economic environment, and 2) characterization of the economic environment. To perform these well, the description is a job of science, and the characterization is a job of art. The more clear and accurate the description of the economic environment is, the higher the possibility is of the correctness of the theoretical conclusions. The more refined the characterization of the economic environment is, the simpler and easier the arguments and conclusions will obtain.

Modern economics provides various perspectives or angles to look at real world economic issues. An economic phenomenon or issue may be very complicated and be affected by many factors. The approach of characterizing the economic environment can grasp the most essential factors of the issue and take our attention to the most key and core characteristics of an issue so that we can avoid unimportant details. An economic environment



usually consists of (1) a number of individuals, (2) the individuals' characteristics, such as preferences, technologies, endowments, etc. (3) informational structures, and (4) institutional economic environments that include fundamental rules for establishing the basis for production, exchange, and distribution.

## **2. Imposition of Behavior Assumptions**

The second step for studying an economic issue is to make assumptions on individuals' behavior. Making appropriate assumptions is of fundamental importance for obtaining a valuable economic theory or assessment. A key assumption modern economics makes about an individual's behavior is that an individual is self-interested. This is a main difference between individuals and other subjects. The self-interested behavior assumption is not only reasonable and realistic, but also have a minimum risk. Even this assumption is not suitable to an economic environment, it does not cause a big trouble to the economy even if it is applied to the economy. A rule of a game designed for self-interested individuals is likely also suitable for altruists, but the reverse is likely not true.

## **3. Adoption of Economic Institutional Arrangement**

The third step for studying an economic issue is to adopt the economic institutional arrangements, which are also called economic mechanisms, which can be regarded as the rules of the game. Depending on the problem under consideration, an economic institutional arrangement could be exogenously given or endogenously determined. For instance, when studying individuals' decisions in the theories of consumers and producers, one implicitly assumes that the undertaken mechanism is a competitive market mechanism takes it as given. However, when considering the choice of economic institutions and arguing the optimality of the market mechanism, the market institution is endogenously determined. The alternative mechanisms that are designed to solve the problem of market failure are also endogenously determined. Economic arrangements should be designed differently for different economic environments and behavior assumptions.

## 4. Determination of Equilibria

The fourth step for studying an economic issue is to make trade-off choices and determine the "best" one. Once given an economic environment, institutional arrangement, and other constraints, such as technical, resource, and budget constraints, individuals will react, based on their incentives and own behavior, and choose an outcome from among the available or feasible outcomes. Such a state is called *equilibrium* and the outcome an *equilibrium outcome*. This is the most general definition an economic "equilibrium".

## 5. Evaluations

The fifth step in studying an economic issue is to evaluate outcomes resulting from the undertaken institutional arrangement and to make value judgments of the chosen equilibrium outcome and economic mechanism based on certain criterion. The most important criterion adopted in modern economics is the notion of efficiency or the "first best". If an outcome is not efficient, there is room for improvement. The other criteria include equity, fairness, incentive-compatibility, informational efficiency, and operation costs for running an economic mechanism.

In summary, in studying an economic issue, one should start by specifying economic environments and then study how individuals interact under the self-interested motion of the individuals within an exogenously given or endogenously determined mechanism. Economists usually use "equilibrium," "efficiency", "information", and "incentive-compatibility" as focal points, and investigate the effects of various economic mechanisms on the behavior of agents and economic units, show how individuals reach equilibria, and evaluate the status at equilibrium. Analyzing an economic problem using such a basic analytical framework has not only consistence in methodology, but also in getting surprising (but logically consistent) conclusions.

### 1.1.4 Methodologies for Studying Modern Economics

As discussed above, any economic theory usually consists of five aspects. Discussions on these five steps will naturally amplify into how to combine these five aspects organically.

To do so, economists usually integrate various studying methods into their analysis. Two methods used in modern economics are providing various levels and aspects studying platforms and establishing reference/benchmark systems.

### **Studying Platform**

A studying platform in modern economics consists of some basic economic theories or principles. It provides a basis for extending the existing theories and analyzing more deep economic issues. Examples of studying platforms are:

- (1) Consumer and producer theories provide a bedrock platform for studying individuals' independent decision choices.
- (2) The general equilibrium theory is based on the theories of consumers and producers and is a higher level platform. It provides a basis for studying interactions of individuals within a market institution and how the market equilibrium is reached in each market.
- (3) The mechanism design theory provides an even higher level of studying platform and can be used to study or design an economic institution. It can be used to compare various economic institutions or mechanisms, as well as to identify which one may be an "optima".

### **Reference Systems/Benchmark**

Modern economics provides various reference/benchmark systems for comparison and to see how far a real world is from an ideal status. A reference system is a standard economic model/theory that results in desired or ideal results, such as efficiency/the "first best". The importance of a reference system does not rely on whether or not it describes the real world correctly or precisely, but instead gives a criterion for understanding the real world. It is a mirror that lets us see the distance between various theoretical models/realistic economic mechanisms and the one given by the reference system. For instance, the general equilibrium theory we will study in the notes is such a reference system. With this reference system, we can study and compare equilibrium outcomes under various market structures with the ideal case of the perfectly competitive mechanism. The first-best

results in a complete information economic environment in information economics. Other examples include the Coase Theorem in property rights theory and economic law, and the Modigliani-Miller Theorem in corporate finance theory.

Although those economic theories or economic models as reference systems may impose some unrealistic assumptions, they are still very useful, and can be used to make further analysis. They establish criteria to evaluate various theoretical models or economic mechanisms used in the real world. A reference system is not required, in most cases it is actually not needed, to predicate the real world well, but it is used to provide a benchmark to see how far a reality is from the ideal status given by a reference system. The value of a reference system is not that it can directly explain the world, but that it provides a benchmark for developing new theories to explain the world. In fact, the establishment of a reference system is very important for any scientific subject, including modern economics. Anyone can talk about an economic issue but the main difference is that a person with systematic training in modern economics has a few reference systems in her mind while a person without training in modern economics does not so he cannot grasp essential parts of the issue and cannot provide deep analysis and insights.

## **Analytical Tools**

Modern economics also provides various powerful analytical tools that are usually given by geometrical or mathematical models. Advantages of such tools can help us to analyze complicated economic behavior and phenomena through a simple diagram or mathematical structure in a model. Examples include (1) the demand-supply curve model, (2) Samuelson's overlapping generation model, (3) the principal-agent model, and (4) the game theoretical model.

### **1.1.5 Roles, Generality, and Limitation of Economic Theory**

#### **Roles of Economic Theory**

An economic theory has three possible roles: (1) It can be used to explain economic behavior and economic phenomena in the real world. (2) It can make scientific predictions or deductions about possible outcomes and consequences of adopted economic mechanisms when economic environments and individuals' behavior are appropriately described. (3)

It can be used to refute faulty goals or projects before they are actually undertaken. If a conclusion is not possible in theory, then it is not possible in a real world setting, as long as the assumptions were approximated realistically.

### **Generality of Economic Theory**

An economic theory is based on assumptions imposed on economic environments, individuals' behavior, and economic institutions. The more general these assumptions are, the more powerful, useful, or meaningful the theory that comes from them is. The general equilibrium theory is considered such a theory.

### **Limitation of Economic Theory**

When examining the generality of an economic theory, one should realize any theory or assumption has a boundary, limitation, and applicable range of economic theory. Thus, two common misunderstandings in economic theory should be avoided. One misunderstanding is to over-evaluate the role of an economic theory. Every theory is based on some imposed assumptions. Therefore, it is important to keep in mind that every theory is not universal, cannot explain everything, but has its limitation and boundary of suitability. When applying a theory to make an economic conclusion and discuss an economic problem, it is important to notice the boundary, limitation, and applicable range of the theory. It cannot be applied arbitrarily, or a wrong conclusion will be the result.

The other misunderstanding is to under-evaluate the role of an economic theory. Some people consider an economic theory useless because they think assumptions imposed in the theory are unrealistic. In fact, no theory, whether in economics, physics, or any other science, is perfectly correct. The validity of a theory depends on whether or not it succeeds in explaining and predicting the set of phenomena that it is intended to explain and predict. Theories, therefore, are continually tested against observations. As a result of this testing, they are often modified, refined, and even discarded.

The process of testing and refining theories is central to the development of modern economics as a science. One example is the assumption of perfect competition. In reality, no competition is perfect. Real world markets seldom achieve this ideal status. The question is then not whether any particular market is perfectly competitive, almost no

one is. The appropriate question is to what degree models of perfect competition can generate insights about real-world markets. We think this assumption is approximately correct in certain situations. Just like frictionless models in physics, such as in free falling body movement (no air resistance), ideal gas (molecules do not collide), and ideal fluids, frictionless models of perfect competition generate useful insights in the economic world.

It is often heard that someone is claiming they have toppled an existing theory or conclusion, or that it has been overthrown, when some condition or assumption behind it is criticized. This is usually needless claim, because any formal rigorous theory can be criticized at anytime because no assumption can coincide fully with reality or cover everything. So, as long as there are no logic errors or inconsistency in a theory, we cannot say that the theory is wrong. We can only criticize it for being too limited or unrealistic. What economists should do is to weaken or relax the assumptions, and obtain new theories based on old theories. We cannot say though that the new theory topples the old one, but instead that the new theory extends the old theory to cover more general situations and different economic environments.

### **1.1.6 Roles of Mathematics in Modern Economics**

Mathematics has become an important tool in modern economics. Almost every field in modern economics uses mathematics and statistics. The mathematical approach to economic analysis is used when economists make use of mathematical symbols in the statement of a problem and also draw upon known mathematical theorems to aid in reasoning. It is not difficult to understand why the mathematical approach has become a dominant approach since finding the boundary of a theory, developing an analytical framework of a theory, establishing reference systems, and providing analytical tools all need mathematics. If we apply a theoretical result to real world without knowing the boundary of a theory, we may get a very bad consequence and hurt an economy seriously.

Some of the advantages of using mathematics are that (1) the “language” used and the descriptions of assumptions are clearer, more accurate, and more precise, (2) the logical process of analysis is more rigorous and clearly sets the boundaries and limitations of a statement, (3) it can give a new result that may not be easily obtained through observation alone, and (4) it can reduce unnecessary debates and improve or extend existing results.

It should be remarked that, although mathematics is of critical importance in modern economics, economics is not mathematics. Economics uses mathematics as a tool in order to model and analyze various economic problems. Statistics and econometrics are used to test or measure the *accuracy* of our predication, and identify causalities among economic variables.

### **1.1.7 Conversion between Economic and Mathematical Languages**

The result of economics research is an economic conclusion. A valuable economics paper usually consists of three parts: (1) It raises important economic questions and answering these questions become the objectives of a paper. (2) It establishes the economic models and draws and proves the conclusions obtained from the model. (3) It uses non-technical language to explain the results and, if relevant, provides policy suggestions.

Thus, the production of an economic conclusion usually goes in three stages: Stage 1: (non-mathematical language stage) Produce preliminary outputs, propose economic ideas, intuitions, and conjectures. (2) Stage 2: (mathematical language stage) Produce intermediate outputs, give a formal and rigorous result through mathematical modeling that specifies the boundary of a theory. Stage 3: (non-technical language stage) Produce final outputs, conclusions, insights, and statements that can be understood by non-specialists.

### **1.1.8 Distinguish between Necessary and Sufficient Conditions as well Normative and Positive Statements**

When discussing an economic issue, it is very important to distinguish between: (1) two types of conditions: necessary and sufficient conditions for a statement to be true, and (2) two types of statements: positive analysis and normative analysis. It is easy to confuse the distinction between necessary conditions and sufficient conditions, a problem that results often in incorrect conclusions.

For instance, it is often heard that the market institution should not be used based on the fact that some countries are market economies but remain poor. The reason this logic results in a wrong conclusion is that they did not realize the adoption of a market mechanism is just a necessary condition for a country to be rich, but is not a sufficient

condition. Becoming a rich country also depends on other factors such as political system, social infrastructures, and culture. Additionally, no example of a country can be found so far that it is rich in the long run, that is not a market economy.

A positive statement state facts while normative statement give opinions or value judgments. Distinguishing these two statements can void many unnecessary debates.

## 1.2 Language and Methods of Mathematics

This section reviews some basic mathematics results such as: continuity and concavity of functions, Separating Hyperplane Theorem, optimization, correspondences (point to set mappings), fixed point theorems, KKM lemma, maximum theorem, etc, which will be used to prove some results in the lecture notes. For good references about the materials discussed in this section, see appendixes in Hildenbrand and Kirman (1988), Mas-Colell (1995), and Varian (1992).

### 1.2.1 Functions

Let  $X$  and  $Y$  be two subsets of Euclidian spaces. In this text, vector inequalities,  $\geq$ ,  $\geq$ , and  $>$ , are defined as follows: Let  $a, b \in \mathbb{R}^n$ . Then  $a \geq b$  means  $a_s \geq b_s$  for all  $s = 1, \dots, n$ ;  $a \geq b$  means  $a \geq b$  but  $a \neq b$ ;  $a > b$  means  $a_s > b_s$  for all  $s = 1, \dots, n$ .

**Definition 1.2.1** A function  $f : X \rightarrow \mathbb{R}$  is said to be *continuous* if at point  $x_0 \in X$ ,

$$\lim_{x \rightarrow x_0} f(x) = f(x_0),$$

or equivalently, for any  $\epsilon > 0$ , there is a  $\delta > 0$  such that for any  $x \in X$  satisfying  $|x - x_0| < \delta$ , we have

$$|f(x) - f(x_0)| < \epsilon$$

A function  $f : X \rightarrow \mathbb{R}$  is said to be continuous on  $X$  if  $f$  is continuous at every point  $x \in X$ .

The idea of continuity is pretty straightforward: There is no disconnected point if we draw a function as a curve. A function is continuous if “small” changes in  $x$  produces “small” changes in  $f(x)$ .



The so-called upper semi-continuity and lower semi-continuity continuities are weaker than continuity. Even weak conditions on continuity are transfer continuity which characterize many optimization problems and can be found in Tian (1992, 1993, 1994) and Tian and Zhou (1995), and Zhou and Tian (1992).

**Definition 1.2.2** A function  $f : X \rightarrow \mathbb{R}$  is said to be *upper semi-continuous* if at point  $x_0 \in X$ , we have

$$\limsup_{x \rightarrow x_0} f(x) \leq f(x_0),$$

or equivalently, for any  $\epsilon > 0$ , there is a  $\delta > 0$  such that for any  $x \in X$  satisfying  $|x - x_0| < \delta$ , we have

$$f(x) < f(x_0) + \epsilon.$$

Although all the three definitions on the upper semi-continuity at  $x_0$  are equivalent, the second one is easier to be versified.

A function  $f : X \rightarrow \mathbb{R}$  is said to be *upper semi-continuous* on  $X$  if  $f$  is upper semi-continuous at every point  $x \in X$ .

**Definition 1.2.3** A function  $f : X \rightarrow \mathbb{R}$  is said to be *lower semi-continuous* on  $X$  if  $-f$  is upper semi-continuous.

It is clear that a function  $f : X \rightarrow \mathbb{R}$  is continuous on  $X$  if and only if it is both upper and lower semi-continuous, or equivalently, for all  $x \in X$ , the upper contour set  $U(x) \equiv \{x' \in X : f(x') \geq f(x)\}$  and the lower contour set  $L(x) \equiv \{x' \in X : f(x') \leq f(x)\}$  are closed subsets of  $X$ .

Let  $f$  be a function on  $\mathbb{R}^k$  with continuous partial derivatives. We define the gradient of  $f$  to be the vector

$$Df(x) = \left[ \frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_k} \right].$$

Suppose  $f$  has continuous second order partial derivatives. We define the Hessian of  $f$  at  $x$  to be the  $n \times n$  matrix denoted by  $D^2f(x)$  as

$$D^2f(x) = \left[ \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right],$$

which is symmetric since

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}.$$

**Definition 1.2.4** A function  $f : X \rightarrow \mathbb{R}$  is said to be *homogeneous of degree  $k$*  if  $f(tx) = t^k f(x)$

An important result concerning homogeneous function is the following:

**Theorem 1.2.1 (Euler's Theorem)** *If a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is homogeneous of degree  $k$  if and only if*

$$kf(x) = \sum_{i=1}^n \frac{\partial f(x)}{\partial x_i} x_i.$$

## 1.2.2 Separating Hyperplane Theorem

A set  $X \subset \mathbb{R}^n$  is said to be **compact** if it is bounded and closed. A set  $X$  is said to be **convex** if for any two points  $x, x' \in X$ , the point  $tx + (1 - t)x' \in X$  for all  $0 \leq t \leq 1$ . Geometrically the convex set means every point on the line segment joining any two points in the set is also in the set.

**Theorem 1.2.2 (Separating Hyperplane Theorem)** *Suppose that  $A, B \subset \mathbb{R}^m$  are convex and  $A \cap B = \emptyset$ . Then, there is a vector  $p \in \mathbb{R}^m$  with  $p \neq 0$ , and a value  $c \in \mathbb{R}$  such that*

$$px \leq c \leq py \quad \forall x \in A \text{ \& } y \in B.$$

*Furthermore, suppose that  $B \subset \mathbb{R}^m$  is convex and closed,  $A \subset \mathbb{R}^m$  is convex and compact, and  $A \cap B = \emptyset$ . Then, there is a vector  $p \in \mathbb{R}^m$  with  $p \neq 0$ , and a value  $c \in \mathbb{R}$  such that*

$$px < c < py \quad \forall x \in A \text{ \& } y \in B.$$

## 1.2.3 Concave and Convex Functions

Concave, convex, and quasi-concave functions arise frequently in microeconomics and have strong economic meanings. They also have a special role in optimization problems.

**Definition 1.2.5** Let  $X$  be a convex set. A function  $f : X \rightarrow \mathbb{R}$  is said to be *concave on  $X$*  if for any  $x, x' \in X$  and any  $t$  with  $0 \leq t \leq 1$ , we have

$$f(tx + (1 - t)x') \geq tf(x) + (1 - t)f(x')$$

The function  $f$  is said to be *strictly concave on  $X$*  if

$$f(tx + (1 - t)x') > tf(x) + (1 - t)f(x')$$

for all  $x \neq x' \in X$  and  $0 < t < 1$ .

A function  $f : X \rightarrow \mathbb{R}$  is said to be (strictly) convex on  $X$  if  $-f$  is (strictly) concave on  $X$ .

**Remark 1.2.1** A linear function is both concave and convex. The sum of two concave (convex) functions is a concave (convex) function.

**Remark 1.2.2** When a function  $f$  defined on a convex set  $X$  has continuous second partial derivatives, it is concave (convex) if and only if the Hessian matrix  $D^2f(x)$  is negative (positive) semi-definite on  $X$ . It is strictly concave (strictly convex) if the Hessian matrix  $D^2f(x)$  is negative (positive) definite on  $X$ .

**Remark 1.2.3** The strict concavity of  $f(\mathbf{x})$  can be checked by verifying if the leading principal minors of the Hessian must alternate in sign, i.e.,

$$\begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} > 0,$$

$$\begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} < 0,$$

and so on, where  $f_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ . This algebraic condition is useful for checking second-order conditions.

In economic theory quasi-concave functions are used frequently, especially for the representation of utility functions. Quasi-concave is somewhat weaker than concavity.

**Definition 1.2.6** Let  $X$  be a convex set. A function  $f : X \rightarrow \mathbb{R}$  is said to be *quasi-concave on  $X$*  if the set

$$\{x \in X : f(x) \geq c\}$$

is convex for all real numbers  $c$ . It is *strictly quasi-concave on  $X$*  if

$$\{x \in X : f(x) > c\}$$

is convex for all real numbers  $c$ .

A function  $f : X \rightarrow \mathbb{R}$  is said to be (strictly) *quasi-convex on  $X$*  if  $-f$  is (strictly) quasi-concave on  $X$ .

**Remark 1.2.4** The sum of two quasi-concave functions in general is not a quasi-concave function. Any monotonic function defined on a subset of the one dimensional real space is both quasi-concave and quasi-convex.

**Remark 1.2.5** When a function  $f$  defined on a convex set  $X$  has continuous second partial derivatives, it is strictly quasi-concave (convex) if the naturally ordered principal minors of the bordered Hessian matrix  $\bar{H}(x)$  alternate in sign, i.e.,

$$\begin{vmatrix} 0 & f_1 & f_2 \\ f_1 & f_{11} & f_{12} \\ f_2 & f_{21} & f_{22} \end{vmatrix} > 0,$$

$$\begin{vmatrix} 0 & f_1 & f_2 & f_3 \\ f_1 & f_{11} & f_{12} & f_{13} \\ f_2 & f_{21} & f_{22} & f_{23} \\ f_3 & f_{31} & f_{32} & f_{33} \end{vmatrix} < 0,$$

and so on.

## 1.2.4 Optimization

Optimization is a fundamental tool for the development of modern microeconomics analysis. Most economic models are based on the the solution of optimization problems. Results of this subsection are used throughout the text.

The basic optimization problem is that of maximizing or minimizing a function on some set. The basic and central result is the existence theorem of Weierstrass.

**Theorem 1.2.3 (Weierstrass Theorem)** *Any upper (lower) semi continuous function reaches its maximum (minimum) on a compact set, and the set of maximum is compact.*

## EQUALITY CONSTRAINED OPTIMIZATION

An optimization problem with equality constraints has the form

$$\begin{aligned} & \max f(\mathbf{x}) \\ \text{such that} \quad & h_1(\mathbf{x}) = d_1 \\ & h_2(\mathbf{x}) = d_2 \\ & \vdots \\ & h_k(\mathbf{x}) = d_k, \end{aligned}$$

where  $f, h_1, \dots, h_k$  are differentiable functions defined on  $\mathbb{R}^n$  and  $k < n$  and  $d_1, \dots, d_k$  are constants.

The most important result for constrained optimization problems is the Lagrange multiplier theorem, giving necessary conditions for a point to be a solution.

Define the Lagrange function:

$$L(x, \lambda) = f(x) + \sum_{i=1}^k \lambda_i [d_i - h_i(x)],$$

where  $\lambda_1, \dots, \lambda_k$  are called the Lagrange multipliers.

The necessary conditions for  $x$  to solve the maximization problem is that there are  $\lambda_1, \dots, \lambda_k$  such that the first-order conditions (FOC) are held:

$$\frac{L(x, \lambda)}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} - \sum_{l=1}^k \lambda_l \frac{\partial h_l(x)}{\partial x_i} = 0 \quad i = 1, 2, \dots, n.$$

## INEQUALITY CONSTRAINED OPTIMIZATION

Consider an optimization problem with inequality constraints:

$$\begin{aligned} & \max f(\mathbf{x}) \\ \text{such that} \quad & g_i(\mathbf{x}) \leq d_i \quad i = 1, 2, \dots, k. \end{aligned}$$

A point  $x$  making all constraints held with equality (i.e.,  $g_i(x) = d_i$  for all  $i$ ) is said to satisfy the constrained qualification condition if the gradient vectors,  $Dg_1(x), Dg_2(x), \dots, Dg_k(x)$  are linearly independent.

**Theorem 1.2.4 (Kuhn-Tucker Theorem)** *Suppose  $x$  solves the inequality constrained optimization problem and satisfies the constrained qualification condition. Then, there are a set of Kuhn-Tucker multipliers  $\lambda_i \geq 0$ ,  $i = 1, \dots, k$ , such that*

$$Df(x) = \sum_{i=1}^k \lambda_i Dg_i(x).$$

Furthermore, we have the complementary slackness conditions:

$$\begin{aligned}\lambda_i &\geq 0 && \text{for all } i = 1, 2, \dots, k \\ \lambda_i &= 0 && \text{if } g_i(x) < D_i.\end{aligned}$$

Comparing the Kuhn-Tucker theorem to the Lagrange multipliers in the equality constrained optimization problem, we see that the major difference is that the signs of the Kuhn-Tucker multipliers are nonnegative while the signs of the Lagrange multipliers can be anything. This additional information can occasionally be very useful.

The Kuhn-Tucker theorem only provides a necessary condition for a maximum. The following theorem states conditions that guarantee the above first-order conditions are sufficient.

**Theorem 1.2.5 (Kuhn-Tucker Sufficiency)** *Suppose  $f$  is concave and each  $g_i$  is convex. If  $x$  satisfies the Kuhn-Tucker first-order conditions specified in the above theorem, then  $x$  is a global solution to the constrained optimization problem.*

We can weaken the conditions in the above theorem when there is only one constraint. Let  $C = \{x \in \mathbb{R}^n : g(x) \leq d\}$ .

**Proposition 1.2.1** *Suppose  $f$  is quasi-concave and the set  $C$  is convex (this is true if  $g$  is quasi-convex). If  $x$  satisfies the Kuhn-Tucker first-order conditions, then  $x$  is a global solution to the constrained optimization problem.*

Sometimes we require  $x$  to be nonnegative. Suppose we had optimization problem:

$$\begin{aligned}\max & f(\mathbf{x}) \\ \text{such that} & \quad g_i(\mathbf{x}) \leq d_i \quad i = 1, 2, \dots, k \\ & x \geq 0.\end{aligned}$$

Then the Lagrange function in this case is given by

$$L(x, \lambda) = f(x) + \sum_{l=1}^k \lambda_l [d_l - h_l(x)] + \sum_{j=1}^n \mu_j x_j,$$

where  $\mu_1, \dots, \mu_k$  are the multipliers associated with constraints  $x_j \geq 0$ . The first-order conditions are

$$\begin{aligned}\frac{L(x, \lambda)}{\partial x_i} &= \frac{\partial f(x)}{\partial x_i} - \sum_{l=1}^k \lambda_l \frac{\partial g_l(x)}{\partial x_i} + \mu_i = 0 \quad i = 1, 2, \dots, n \\ \lambda_l &\geq 0 \quad l = 1, 2, \dots, k \\ \lambda_l &= 0 \quad \text{if } g_l(x) < d_l \\ \mu_i &\geq 0 \quad i = 1, 2, \dots, n \\ \mu_i &= 0 \quad \text{if } x_i > 0.\end{aligned}$$

Eliminating  $\mu_i$ , we can equivalently write the above first-order conditions with nonnegative choice variables as

$$\frac{L(x, \lambda)}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} - \sum_{l=1}^k \lambda_l \frac{\partial g_l(x)}{\partial x_i} \leq 0 \quad \text{with equality if } x_i > 0 \quad i = 1, 2, \dots, n,$$

or in matrix notation,

$$Df - \lambda Dg \leq 0$$

$$x[Df - \lambda Dg] = 0$$

where we have written the product of two vector  $x$  and  $y$  as the inner production, i.e.,  $xy = \sum_{i=1}^n x_i y_i$ . Thus, if we are at an interior optimum (i.e.,  $x_i > 0$  for all  $i$ ), we have

$$Df(x) = \lambda Dg.$$

### 1.2.5 The Envelope Theorem

Consider an arbitrary maximization problem where the objective function depends on some parameter  $a$ :

$$M(a) = \max_x f(x, a).$$

The function  $M(a)$  gives the maximized value of the objective function as a function of the parameter  $a$ .

Let  $x(a)$  be the value of  $x$  that solves the maximization problem. Then we can also write  $M(a) = f(x(a), a)$ . It is often of interest to know how  $M(a)$  changes as  $a$  changes.

The envelope theorem tells us the answer:

$$\frac{dM(a)}{da} = \left. \frac{\partial f(x, a)}{\partial a} \right|_{x=x(a)}.$$

This expression says that the derivative of  $M$  with respect to  $a$  is given by the partial derivative of  $f$  with respect to  $a$ , holding  $x$  fixed at the optimal choice. This is the meaning of the vertical bar to the right of the derivative. The proof of the envelope theorem is a relatively straightforward calculation.

Now consider a more general parameterized constrained maximization problem of the form

$$\begin{aligned} M(a) &= \max_{x_1, x_2} g(x_1, x_2, a) \\ \text{such that } h(x_1, x_2, a) &= 0. \end{aligned}$$

The Lagrangian for this problem is

$$\mathcal{L} = g(x_1, x_2, a) - \lambda h(x_1, x_2, a),$$

and the first-order conditions are

$$\begin{aligned} \frac{\partial g}{\partial x_1} - \lambda \frac{\partial h}{\partial x_1} &= 0 \\ \frac{\partial g}{\partial x_2} - \lambda \frac{\partial h}{\partial x_2} &= 0 \\ h(x_1, x_2, a) &= 0. \end{aligned} \tag{1.1}$$

These conditions determine the optimal choice functions  $(x_1(a), x_2(a), a)$ , which in turn determine the maximum value function

$$M(a) \equiv g(x_1(a), x_2(a)). \tag{1.2}$$

The envelope theorem gives us a formula for the derivative of the value function with respect to a parameter in the maximization problem. Specifically, the formula is

$$\begin{aligned} \frac{dM(a)}{da} &= \left. \frac{\partial \mathcal{L}(x, a)}{\partial a} \right|_{x=x(a)} \\ &= \left. \frac{\partial g(x_1, x_2, a)}{\partial a} \right|_{x_i=x_i(a)} - \lambda \left. \frac{\partial h(x_1, x_2, a)}{\partial a} \right|_{x_i=x_i(a)} \end{aligned}$$



As before, the interpretation of the partial derivatives needs special care: they are the derivatives of  $g$  and  $h$  with respect to  $a$  holding  $x_1$  and  $x_2$  fixed at their optimal values.

### 1.2.6 Point-to-Set Mappings

When a mapping is not a single-valued function, but is a point-to-set mapping, it is called a correspondence, or multi-valued functions. That is, a correspondence  $F$  maps point  $x$  in the domain  $X \subseteq \mathbb{R}^n$  into sets in the range  $Y \subseteq \mathbb{R}^m$ , and it is denoted by  $F : X \rightarrow 2^Y$ . We also use  $F : X \rightarrow \rightarrow Y$  to denote the mapping  $F : X \rightarrow 2^Y$  in this lecture notes.

**Definition 1.2.7** A correspondence  $F : X \rightarrow 2^Y$  is: (1) *non-empty valued* if the set  $F(x)$  is non-empty for all  $x \in X$ ; (2) *convex valued* if the set  $F(x)$  is a convex set for all  $x \in X$ ; (3) *closed valued* if the set  $F(x)$  is a closed set for all  $x \in X$ ; (4) *compact valued* if the set  $F(x)$  is a compact set for all  $x \in X$ .

Intuitively, a correspondence is continuous if small changes in  $x$  produce small changes in the set  $F(x)$ . Unfortunately, giving a formal definition of continuity for correspondences is not so simple. Figure 1.1 shows a continuous correspondence.

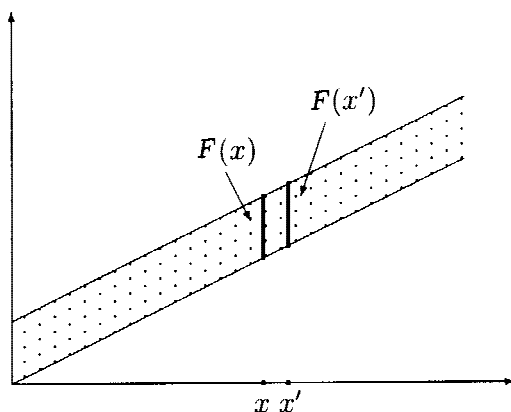


Figure 1.1: A Continuous correspondence.

The notions of hemi-continuity are usually defined in terms of sequences (see Debreu (1959) and Mask-Collell et al. (1995)), but, although they are relatively easy to verify, they

are not intuitive and depend on the assumption that a correspondence is compact-valued. The following definitions are more formal (see, Border, 1988).

**Definition 1.2.8** A correspondence  $F : X \rightarrow 2^Y$  is *upper hemi-continuous* at  $x$  if for each open set  $U$  containing  $F(x)$ , there is an open set  $N(x)$  containing  $x$  such that if  $x' \in N(x)$ , then  $F(x') \subset U$ . A correspondence  $F : X \rightarrow 2^Y$  is *upper hemi-continuous* if it is upper hemi-continuous at every  $x \in X$ , or equivalently, if the set  $\{x \in X : F(x) \subset V\}$  is open in  $X$  for every open set subset  $V$  of  $Y$ .

**Remark 1.2.6** Upper hemi-continuity captures the idea that  $F(x)$  will not “suddenly contain new points” just as we move past some point  $x$ , in other words,  $F(x)$  does not suddenly becomes much larger if one changes the argument  $x$  slightly. That is, if one starts at a point  $x$  and moves a little way to  $x'$ , upper hemi-continuity at  $x$  implies that there will be no point in  $F(x')$  that is not close to some point in  $F(x)$ .

**Definition 1.2.9** A correspondence  $F : X \rightarrow 2^Y$  is said to be *lower hemi-continuous at  $x$*  if for every open set  $V$  with  $F(x) \cap V \neq \emptyset$ , there exists a neighborhood  $N(x)$  of  $x$  such that  $F(x') \cap V \neq \emptyset$  for all  $x' \in N(x)$ . A correspondence  $F : X \rightarrow 2^Y$  is *lower hemi-continuous* if it is lower hemi-continuous at every  $x \in X$ , or equivalently, the set  $\{x \in X : F(x) \cap V \neq \emptyset\}$  is open in  $X$  for every open set  $V$  of  $Y$ .

**Remark 1.2.7** Lower hemi-continuity captures the idea that any element in  $F(x)$  can be “approached” from all directions, in other words,  $F(x)$  does not suddenly becomes much smaller if one changes the argument  $x$  slightly. That is, if one starts at some point  $x$  and some point  $y \in F(x)$ , lower hemi-continuity at  $x$  implies that if one moves a little way from  $x$  to  $x'$ , there will be some  $y' \in F(x')$  that is close to  $y$ .

**Remark 1.2.8** Based on the following two facts, both notions of hemi-continuity can be characterized by sequences.

- (a) If a correspondence  $F : X \rightarrow 2^Y$  is compact-valued, then it is upper hemi-continuous if and only if for any  $\{x_k\}$  with  $x_k \rightarrow x$  and  $\{y_k\}$  with  $y_k \in F(x_k)$ , there exists a converging subsequence  $\{y_{k_m}\}$  of  $\{y_k\}$ ,  $y_{k_m} \rightarrow y$ , such that  $y \in F(x)$ .

- (b) A correspondence  $F : X \rightarrow 2^Y$  is said to be *lower hemi-continuous at  $x$*  if and only if for any  $\{x_k\}$  with  $x_k \rightarrow x$  and  $y \in F(x)$ , then there is a sequence  $\{y_k\}$  with  $y_k \rightarrow y$  and  $y_k \in F(x_k)$ .

**Definition 1.2.10** A correspondence  $F : X \rightarrow 2^Y$  is said to be *closed at  $x$*  if for any  $\{x_k\}$  with  $x_k \rightarrow x$  and  $\{y_k\}$  with  $y_k \rightarrow y$  and  $y_k \in F(x_k)$  implies  $y \in F(x)$ .  $F$  is said to be closed if  $F$  is closed for all  $x \in X$  or equivalently

$$Gr(F) = \{(x, y) \in X \times Y : y \in F(x)\} \text{ is closed.}$$

**Remark 1.2.9** Regarding the relationship between upper hemi-continuity and closed graph, the following facts can be proved.

- (i) If  $Y$  is compact and  $F : X \rightarrow 2^Y$  is closed-valued, then  $F$  has closed graph implies it is upper hemi-continuous.
- (ii) If  $X$  and  $Y$  are closed and  $F : X \rightarrow 2^Y$  is closed-valued, then  $F$  is upper hemi-continuous implies that it has closed graph.

Because of fact (i), a correspondence with closed graph is sometimes called upper hemi-continuity in the literature. But one should keep in mind that they are not the same in general. For example, let  $F : \mathbb{R}_+ \rightarrow 2^{\mathbb{R}}$  be defined by

$$F(x) = \begin{cases} \{\frac{1}{x}\} & \text{if } x > 0 \\ \{0\} & \text{if } x = 0. \end{cases}$$

The correspondence is closed but not upper hemi-continuous.

Also, define  $F : \mathbb{R}_+ \rightarrow 2^{\mathbb{R}}$  by  $F(x) = (0, 1)$ . Then  $F$  is upper hemi-continuous but not closed.

Figure 1.2 shows the correspondence is upper hemi-continuous, but not lower hemi-continuous. To see why it is upper hemi-continuous, imagine an open interval  $U$  that encompasses  $F(x)$ . Now consider moving a little to the left of  $x$  to a point  $x'$ . Clearly  $F(x') = \{\hat{y}\}$  is in the interval. Similarly, if we move to a point  $x'$  a little to the right of  $x$ , then  $F(x)$  will be inside the interval so long as  $x'$  is sufficiently close to  $x$ . So it is upper hemi-continuous. On the other hand, the correspondence is not lower hemi-continuous. To see this, consider the point  $y \in F(x)$ , and let  $U$  be a very small interval around  $y$  that

does not include  $\hat{y}$ . If we take any open set  $N(x)$  containing  $x$ , then it will contain some point  $x'$  to the left of  $x$ . But then  $F(x') = \{\hat{y}\}$  will contain no points near  $y$ , i.e., it will not intersect  $U$ .

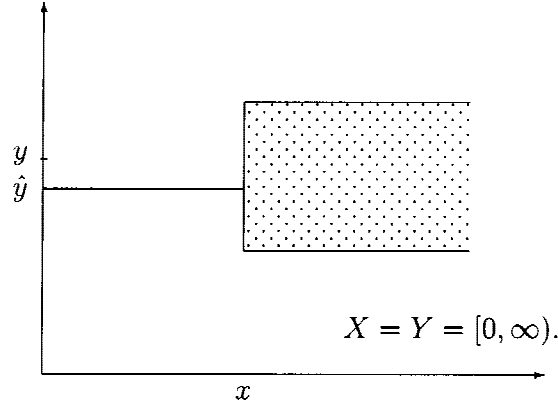


Figure 1.2: A correspondence that is upper hemi-continuous, but not lower hemi-continuous.

Figure 1.3 shows the correspondence is lower hemi-continuous, but not upper hemi-continuous. To see why it is lower hemi-continuous. For any  $0 \leq x' \leq x$ , note that  $F(x') = \{\hat{y}\}$ . Let  $x_n = x' - 1/n$  and let  $y_n = \hat{y}$ . Then  $x_n > 0$  for sufficiently large  $n$ ,  $x_n \rightarrow x'$ ,  $y_n \rightarrow \hat{y}$ , and  $y_n \in F(x_n) = \{\hat{y}\}$ . So it is lower hemi-continuous. It is clearly lower hemi-continuous for  $x_i > x$ . Thus, it is lower hemi-continuous on  $X$ . On the other hand, the correspondence is not upper hemi-continuous. If we start at  $x$  by noting that  $F(x) = \{\hat{y}\}$ , and make a small move to the right to a point  $x'$ , then  $F(x')$  suddenly contains many points that are not close to  $\hat{y}$ . So this correspondence fails to be upper hemi-continuous.

Combining upper and lower hemi-continuity, we can define the continuity of a correspondence.

**Definition 1.2.11** A correspondence  $F : X \rightarrow 2^Y$  at  $x \in X$  is said to be *continuous* if it is both upper hemi-continuous and lower hemi-continuous at  $x \in X$ . A correspondence  $F : X \rightarrow 2^Y$  is said to be *continuous* if it is both upper hemi-continuous and lower hemi-continuous.

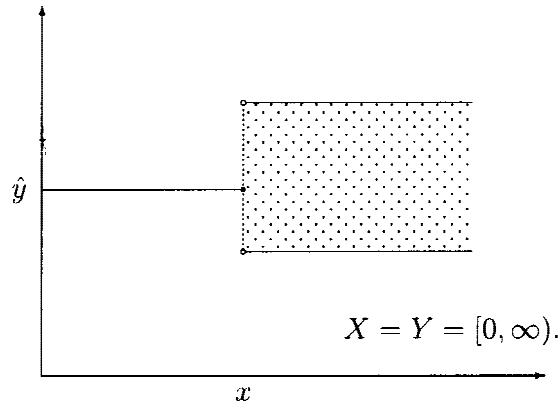


Figure 1.3: A correspondence that is lower hemi-continuous, but not upper hemi-continuous.

**Remark 1.2.10** As it turns out, the notions of upper and lower hemi-continuous correspondence both reduce to the standard notion of continuity for a function if  $F(\cdot)$  is a single-valued correspondence, i.e., a function. That is,  $F(\cdot)$  is a single-valued upper (or lower) hemi-continuous correspondence if and only if it is a continuous function.

**Definition 1.2.12** A correspondence  $F : X \rightarrow 2^Y$  said to be *open* if its graph

$$Gr(F) = \{(x, y) \in X \times Y : y \in F(x)\} \text{ is open.}$$

**Definition 1.2.13** A correspondence  $F : X \rightarrow 2^Y$  said to have *upper open sections* if  $F(x)$  is open for all  $x \in X$ .

A correspondence  $F : X \rightarrow 2^Y$  said to have *lower open sections* if its inverse set  $F^{-1}(y) = \{x \in X : y \in F(x)\}$  is open.

**Remark 1.2.11** If a correspondence  $F : X \rightarrow 2^Y$  has an open graph, then it has upper and lower open sections. If a correspondence  $F : X \rightarrow 2^Y$  has lower open sections, then it must be lower hemi-continuous.

## 1.2.7 Continuity of a Maximum

In many places, we need to check if an optimal solution is continuous in parameters, say, to check the continuity of the demand function. We can apply the so-called Maximum Theorem.

**Theorem 1.2.6 (Berg’s Maximum Theorem)** Suppose  $f(x, a)$  is a continuous function mapping from  $A \times X \rightarrow \mathbb{R}$ , and the constraint set  $F : A \rightarrow \rightarrow X$  is a continuous correspondence with non-empty compact values. Then, the optimal valued function (also called marginal function):

$$M(a) = \max_{x \in F(a)} f(x, a)$$

is a continuous function, and the optimal solution:

$$\phi(a) = \arg \max_{x \in F(a)} f(x, a)$$

is a upper hemi-continuous correspondence.

## 1.2.8 Fixed Point Theorems

To show the existence of a competitive equilibrium for the continuous aggregate excess demand function, we will use the following fixed-point theorem. The generalization of Brouwer’s fixed theorem can be found in Tian (1991) that gives necessary and sufficient conditions for a function to have a fixed point.

**Theorem 1.2.7 (Brouwer’s Fixed Theorem)** Let  $X$  be a non-empty, compact, and convex subset of  $\mathbb{R}^m$ . If a function  $f : X \rightarrow X$  is continuous on  $X$ , then  $f$  has a fixed point, i.e., there is a point  $x^* \in X$  such that  $f(x^*) = x^*$ .

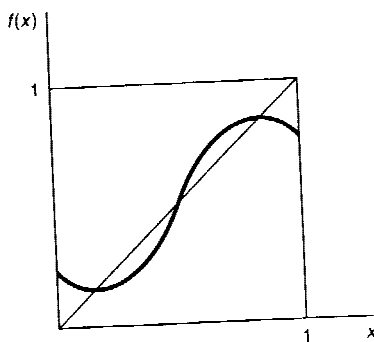


Figure 1.4: Fixed points are given by the intersections of the 45° line and the curve of the function. There are three fixed points in the case depicted.

**Example 1.2.1**  $f : [0, 1] \rightarrow [0, 1]$  is continuous, then  $f$  has a fixed point  $(x)$ . To see this, let

$$g(x) = f(x) - x.$$

Then, we have

$$g(0) = f(0) \geq 0$$

$$g(1) = f(1) - 1 \leq 0.$$

From the mean-value theorem, there is a point  $x^* \in [0, 1]$  such that  $g(x^*) = f(x^*) - x^* = 0$ .

When a mapping is a correspondence, we have the following version of fixed point theorem.

**Theorem 1.2.8 (Kakutani's Fixed Point Theorem)** *Let  $X$  be a non-empty, compact, and convex subset of  $\mathbb{R}^m$ . If a correspondence  $F : X \rightarrow 2^X$  is a upper hemi-continuous correspondence with non-empty compact and convex values on  $X$ , then  $F$  has a fixed point, i.e., there is a point  $x^* \in X$  such that  $x^* \in F(x^*)$ .*

The Knaster-Kuratowski-Mazurkiewicz (KKM) lemma is quite basic and in some ways more useful than Brouwer's fixed point theorem. The following is a generalized version of KKM lemma due to Ky Fan (1984).

**Theorem 1.2.9 (FKKM Theorem)** *Let  $Y$  be a convex set and  $\emptyset \neq X \subset Y$ . Suppose  $F : X \rightarrow 2^Y$  is a correspondence such that*

- (1)  $F(x)$  is closed for all  $x \in X$ ;
- (2)  $F(x_0)$  is compact for some  $x_0 \in X$ ;
- (3)  $F$  is FS-convex, i.e, for any  $x_1, \dots, x_m \in X$  and its convex combination  $x_\lambda = \sum_{i=1}^m \lambda_i x_i$ , we have  $x_\lambda \in \cup_{i=1}^m F(x_i)$ .

*Then  $\cap_{x \in X} F(x) \neq \emptyset$ .*

Here, The term FS is for Fan (1984) and Sonnenschein (1971), who introduced the notion of FS-convexity.

The various characterization results on Kakutani's Fixed Point Theorem, KKM Lemma, and Maximum Theorem can be found in Tian (1991, 1992, 1994) and Tian and Zhou (1992).

# Reference

- Border, K. C., *Fixed Point Theorems with Applications to Economics and Game Theory*, Cambridge: Cambridge University Press, 1985.
- Debreu, G. (1959), *Theory of Value*, (Wiley, New York).
- Fan, K., “Some Properties of Convex Sets Related to Fixed Point Theorem,” *Mathematics Annals*, 266 (1984), 519-537.
- Hildenbrand, W., A. P. Kirman, *Equilibrium Analysis: Variations on the Themes by Edgeworth and Walras*, North-Holland: New York, 1988.
- Jehle, G. A., and P. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, 1998, Chapters 1-2.
- Luenberger, D., *Microeconomic Theory*, McGraw-Hill, 1995, Appendixes A-D.
- Mas Colell, A., M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995, Mathematical Appendix.
- Nessah, R. and G. Tian, “Existence of Solution of Minimax Inequalities, Equilibria in Games and Fixed Points without Convexity and Compactness Assumptions,” (with Rabia Nessah), *Journal of Optimization Theory and Applications*, forthcoming.
- Qian, Y. “Understanding Modern Economics,” *Economic and Social System Comparison*, 2 (2002).
- Rubinstein Ariel, *Lecture Notes in Microeconomics (modeling the economic agent)*, Princeton University Press, 2005.
- Takayama, A. *Mathematical Economics*, the second edition, Cambridge: Cambridge University Press, 1985, Chapters 1-3.
- Tian, G., *The Basic Analytical Framework and Methodologies in Modern Economics*, 2004 (in Chinese). <http://econweb.tamu.edu/tian/chinese.htm>.
- Tian, G., “Fixed Points Theorems for Mappings with Non-Compact and Non-Convex Domains,” *Journal of Mathematical Analysis and Applications*, 158 (1991), 161-167.



- Tian, G. “Generalizations of the FKKM Theorem and Ky-Fan Minimax Inequality, with Applications to Maximal Elements, Price Equilibrium, and Complementarity,” *Journal of Mathematical Analysis and Applications*, 170 (1992), pp. 457-471.
- Tian, G., “Generalized KKM Theorem and Minimax Inequalities and Their Applications,” *Journal of Optimization Theory and Applications*, 83 (1994), 375-389.
- Tian, G., “Minimax Inequalities and Related Theorems for Arbitrary Topological Spaces: A Full Characterization,” 2013, working paper.
- Tian, G. and J. Zhou, “The Maximum Theorem and the Existence of Nash Equilibrium of (Generalized) Games without Lower Semicontinuites,” *Journal of Mathematical Analysis and Applications*, 166 (1992), pp. 351-364.
- Tian, G. and J. Zhou, “Transfer Continuities, Generalizations of the Weierstrass Theorem and Maximum Theorem—A Full Characterization,” *Journal of Mathematical Economics*, 24 (1995), 281-303.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 26-27.

# Part I

## Individual Decision Making

Part I is devoted to the theories of individual decision making and consists of three chapters: consumer theory, producer theory, and choice under uncertainty. It studies how a consumer or producer selects an appropriate action or making an appropriate decision. Microeconomic theory is founded on the premise that these individuals behave rationally, making choices that are optimal for themselves. Throughout this part, we restrict ourselves to an ideal situation (benchmark case) where the behavior of the others are summarized in non-individualized parameters – the prices of commodities, each individual makes decision independently by taking prices as given and individuals' behavior are **indirectly** interacted through prices.

We will treat consumer theory first, and at some length – both because of its intrinsic importance, and because its methods and results are paradigms for many other topic areas. Producer theory is next, and we draw attention to the many formal similarities between these two important building blocks of modern microeconomics. Finally, we conclude our treatment of the individual consumer by looking at the problem of choice under uncertainty. It is greatly acknowledged that the materials are largely drawn from the references in the end of each chapter, especially from Varian (1992).

# Chapter 2

## Consumer Theory

### 2.1 Introduction

In this chapter, we will explore the essential features of modern consumer theory—a bedrock foundation on which so many theoretical structures in economics are build, and it is also central to the economists’ way of thinking.

A consumer can be characterized by many factors and aspects such as sex, age, lifestyle, wealth, parentage, ability, intelligence, etc. But which are most important ones for us to study consumer’s behavior in making choices? To grasp the most important features in studying consumer behavior and choices in modern consumer theory, it is assumed that the key characteristic of a consumer consists of three essential components: the consumption set, initial endowments, and the preference relation. Consumer’s characteristic together with the behavior assumption are building blocks in any model of consumer theory. The consumption set represents the set of all individually feasible alternatives or consumption plans and sometimes also called the choice set. An initial endowment represents the amount of various goods the consumer initially has and can consume or trade with other individuals. The preference relation specifies the consumer’s tastes or satisfactions for the different objects of choice. The behavior assumption expresses the guiding principle the consumer uses to make final choices and identifies the ultimate objects in choice. It is generally assumed that the consumer seeks to identify and select an available alternative that is most preferred in the light of his/her personal tastes/interstes.

## 2.2 Consumption Set and Budget Constraint

### 2.2.1 Consumption Set

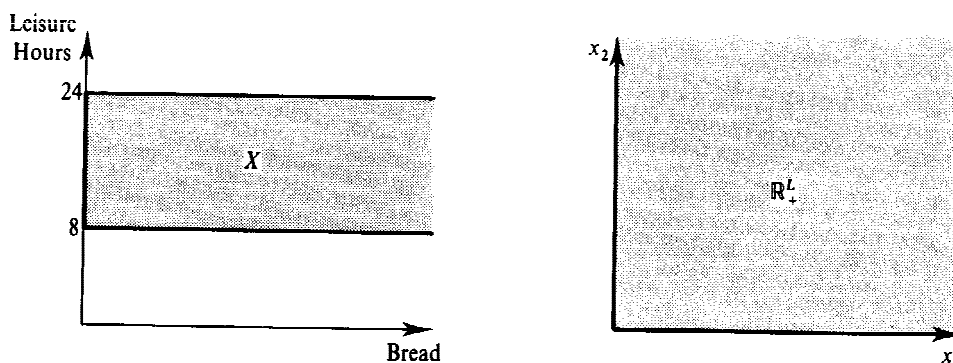


Figure 2.1: The left figure: A consumption set that reflects legal limit on the number of working hours. The right figure: the consumption set  $\mathbb{R}_+^L$ .

We consider a consumer faced with possible consumption bundles in **consumption set**  $X$ . We usually assume that  $X$  is the nonnegative orthant in  $\mathbb{R}^L$  as shown in the right figure in Figure 2.1, but more specific consumption sets may be used. For example, it may allow consumptions of some good in a suitable interval as such leisure as shown in the left figure in Figure 2.1, or we might only include bundles that would give the consumer at least a subsistence existence or that consists of only integer units of consumptions as shown in Figure 2.2.

We assume that  $X$  is a closed and convex set unless otherwise stated. The convexity of a consumption set means that every good is divisible and can be consumed by fraction units.

### 2.2.2 Budget Constraint

In the basic problem of consumer's choice, not all consumptions bundles are affordable in a limited resource economy, and a consumer is constrained by his/her wealth. In a market institution, the wealth may be determined by the value of his/her initial endowment and/or income from stock-holdings of firms. It is assumed that the income or wealth

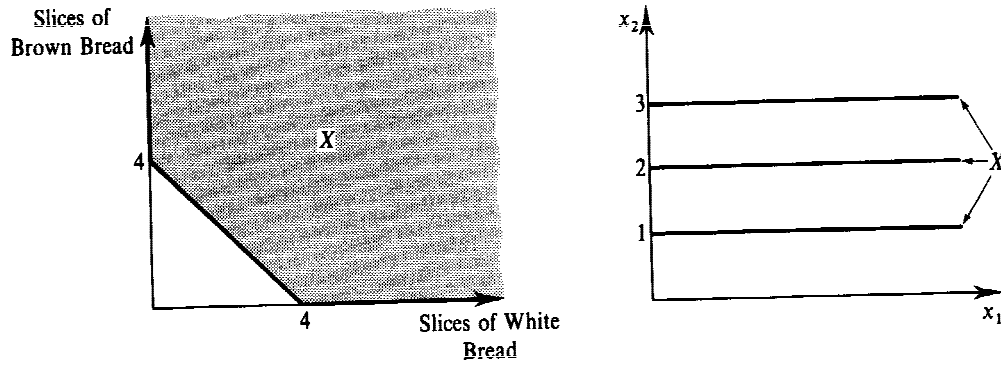


Figure 2.2: The left figure: A consumption set that reflects survival needs. The right figure: A consumption set where good 2 must be consumed in integer amounts.

of the consumer is fixed and the prices of goods cannot be affected by the consumer's consumption when discussing a consumer's choice. Let  $m$  be the fixed amount of money available to a consumer, and let  $\mathbf{p} = (p_1, \dots, p_L)$  be the vector of prices of goods,  $1, \dots, L$ . The set of affordable alternatives is thus just the set of all bundles that satisfy the consumer's budget constraint. The set of affordable bundles, the budget set of the consumer, is given by

$$B(p, m) = \{x \in X : px \leq m, \}$$

where  $px$  is the inner product of the price vector and consumption bundle, i.e.,  $px = \sum_{l=1}^L p_l x_l$  which is the sum of expenditures of commodities at prices  $p$ . The ratio,  $p_l/p_k$ , may be called the **economic rate of substitution** between goods  $i$  and  $j$ . Note that multiplying all prices and income by some positive number does not change the budget set.

Thus, the budget set reflects consumer's objective ability of purchasing commodities and the scarcity of resources. It significantly restricts the consumer choices. To determine the optimal consumption bundles, one needs to combine consumer objective ability of purchasing various commodities with subjective taste on various consumptions bundles which are characterized by the notion of preference or utility.

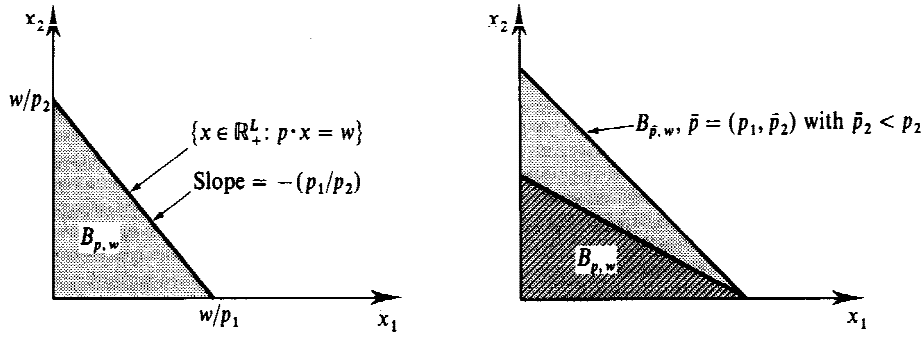


Figure 2.3: The left figure: A budget set. The right figure: The effect of a price change on the budget set.

## 2.3 Preferences and Utility

### 2.3.1 Preferences

The consumer is assumed to have preferences on the consumption bundles in  $X$  so that he can compare and rank various goods available in the economy. When we write  $\mathbf{x} \succeq \mathbf{y}$ , we mean “the consumer thinks that the bundle  $\mathbf{x}$  is at least as good as the bundle  $\mathbf{y}$ .” We want the preferences to order the set of bundles. Therefore, we need to assume that they satisfy the following standard properties.

**COMPLETE.** For all  $\mathbf{x}$  and  $\mathbf{y}$  in  $X$ , either  $\mathbf{x} \succeq \mathbf{y}$  or  $\mathbf{y} \succeq \mathbf{x}$  or both.

**REFLEXIVE.** For all  $\mathbf{x}$  in  $X$ ,  $\mathbf{x} \succeq \mathbf{x}$ .

**TRANSITIVE.** For all  $\mathbf{x}, \mathbf{y}$  and  $\mathbf{z}$  in  $X$ , if  $\mathbf{x} \succeq \mathbf{y}$  and  $\mathbf{y} \succeq \mathbf{z}$ , then  $\mathbf{x} \succeq \mathbf{z}$ .

The first assumption is just says that any two bundles can be compared, the second is trivial and says that every consumption bundle is as good as itself, and the third requires the consumer’s choice be consistent. A preference relation that satisfies these three properties is called a **preference ordering**.

Given an ordering  $\succeq$  describing “weak preference,” we can define the **strict preference**  $\succ$  by  $\mathbf{x} \succ \mathbf{y}$  to mean not  $\mathbf{y} \succeq \mathbf{x}$ . We read  $\mathbf{x} \succ \mathbf{y}$  as “ $\mathbf{x}$  is strictly preferred to  $\mathbf{y}$ .” Similarly, we define a notion of **indifference** by  $\mathbf{x} \sim \mathbf{y}$  if and only if  $\mathbf{x} \succeq \mathbf{y}$  and  $\mathbf{y} \succeq \mathbf{x}$ .

Given a preference ordering, we often display it graphically, as shown in Figure 2.4. The set of all consumption bundles that are indifferent to each other is called an **indiffer-**

**ence curve.** For a two-good case, the slope of an indifference curve at a point measures **marginal rate of substitution** between goods  $x_1$  and  $x_2$ . For a  $L$ -dimensional case, the marginal rate of substitution between two goods is the slope of an indifference surface, measured in a particular direction.

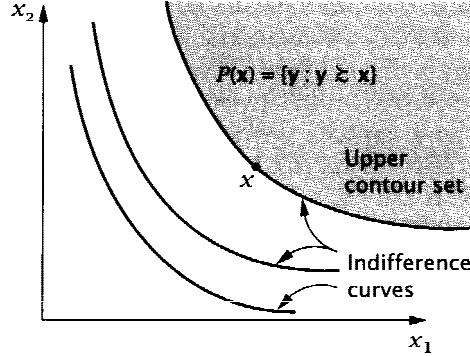


Figure 2.4: Preferences in two dimensions.

For a given consumption bundle  $y$ , let  $P(y) = \{\mathbf{x} \text{ in } X: \mathbf{x} \succeq \mathbf{y}\}$  be the set of all bundles on or above the indifference curve through  $y$  and it is called the **upper contour set** at  $y$ ,  $P_s(y) = \{\mathbf{x} \text{ in } X: \mathbf{x} \succ \mathbf{y}\}$  be the set of all bundles above the indifference curve through  $y$  and it is called the **strictly upper contour set** at  $y$ ,  $L(y) = \{\mathbf{x} \text{ in } X: \mathbf{x} \preceq \mathbf{y}\}$  be the set of all bundles on or below the indifference curve through  $y$  and it is called the **lower contour set** at  $y$ , and  $L_s(y) = \{\mathbf{x} \text{ in } X: \mathbf{x} \prec \mathbf{y}\}$  be the set of all bundles on or below the indifference curve through  $y$  and it is called the **strictly lower contour set** at  $y$ .

We often wish to make other assumptions on consumers' preferences; for example.

**CONTINUITY.** For all  $\mathbf{y}$  in  $X$ , the upper and lower contour sets  $P(y)$  and  $L(y)$ , are closed. It follows that the strictly upper and lower contour sets,  $P_s(y)$  and  $L_s(y)$ , are open sets.

This assumption is necessary to rule out certain discontinuous behavior; it says that if  $(\mathbf{x}^i)$  is a sequence of consumption bundles that are all at least as good as a bundle  $\mathbf{y}$ , and if this sequence converges to some bundle  $\mathbf{x}^*$ , then  $\mathbf{x}^*$  is at least as good as  $\mathbf{y}$ . The most important consequence of continuity is this: if  $\mathbf{y}$  is strictly preferred to  $\mathbf{z}$  and if  $\mathbf{x}$  is a bundle that is close enough to  $\mathbf{y}$ , then  $\mathbf{x}$  must be strictly preferred to  $\mathbf{z}$ .



**Example 2.3.1 (Lexicographic Ordering)** An interesting preference ordering is the so-called lexicographic ordering defined on  $\mathbb{R}^L$ , based on the way one orders words alphabetically. It is defined follows:  $x \succeq y$  if and only if there is a  $l$ ,  $1 \leq l \leq L$ , such that  $x_i = y_i$  for  $i < l$  and  $x_l > y_l$  or if  $x_i = y_i$  for all  $i = 1, \dots, L$ . Essentially the lexicographic ordering compares the components one at a time, beginning with the first, and determines the ordering based on the first time a different component is found, the vector with the greater component is ranked highest. However, the lexicographic ordering is not continuous or even not upper semi-continuous, i.e., the upper contour set is not closed. This is easily seen for the two-dimensional case by considering the upper contour correspondence to  $y = (1, 1)$ , that is, the set  $P(1, 1) = \{x \in X : x \succeq (1, 1)\}$  as shown in Figure 2.5. It is clearly not closed because the boundary of the set below  $(1, 1)$  is not contained in the set.

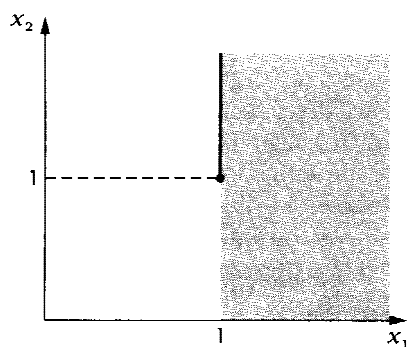


Figure 2.5: Preferred set for lexicographic ordering.

There are two more assumptions, namely, monotonicity and convexity, that are often used to guarantee nice behavior of consumer demand functions. We first give various types of monotonicity properties used in the consumer theory.

**WEAK MONOTONICITY.** *If  $\mathbf{x} \geq \mathbf{y}$  then  $\mathbf{x} \succeq \mathbf{y}$ .*

**MONOTONICITY.** *If  $\mathbf{x} > \mathbf{y}$ , then  $\mathbf{x} \succ \mathbf{y}$ .*

**STRONG MONOTONICITY.** *If  $\mathbf{x} \geq \mathbf{y}$  and  $\mathbf{x} \neq \mathbf{y}$ , then  $\mathbf{x} \succ \mathbf{y}$ .*

Weak monotonicity says that “at least as much of everything is at least as good,” which ensures a commodity is a “good”, but not a “bad”. Monotonicity says that strictly more of every good is strictly better. Strong monotonicity says that at least as much of every good, and strictly more of some good, is strictly better.

Another assumption that is weaker than either kind of monotonicity or strong monotonicity is the following:

**LOCAL NON-SATIATION.** *Given any  $\mathbf{x}$  in  $X$  and any  $\epsilon > 0$ , then there is some bundle  $\mathbf{y}$  in  $X$  with  $|\mathbf{x} - \mathbf{y}| < \epsilon$  such that  $\mathbf{y} \succ \mathbf{x}$ .*

**NON-SATIATION.** *Given any  $\mathbf{x}$  in  $X$ , then there is some bundle  $\mathbf{y}$  in  $X$  such that  $\mathbf{y} \succ \mathbf{x}$ .*

**Remark 2.3.1** Monotonicity of preferences can be interpreted as individuals' desires for goods: the more, the better. Local non-satiation says that one can always do a little bit better, even if one is restricted to only small changes in the consumption bundle. Thus, local non-satiation means individuals' desires are unlimited. You should verify that (strong) monotonicity implies local non-satiation and local non-satiation implies non-satiation, but not vice versa.

We now give various types of convexity properties used in the consumer theory.

**STRICT CONVEXITY.** Given  $x, x'$  in  $X$  such that  $x' \succeq x$ , then it follows that  $tx + (1 - t)x' \succ x$  for all  $0 < t < 1$ .

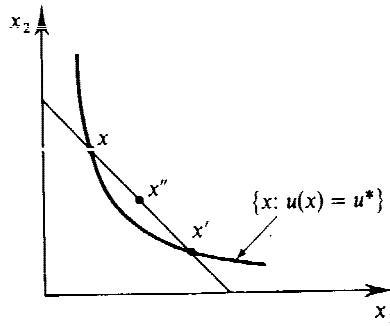


Figure 2.6: Strict convex indifference curves

**CONVEXITY.** Given  $x, x'$  in  $X$  such that  $x' \succ x$ , then it follows that  $tx + (1 - t)x' \succ x$  for all  $0 \leq t < 1$ .

**WEAK CONVEXITY.** Given  $x, x'$  in  $X$  such that  $x' \succeq x$ , then it follows that  $tx + (1 - t)x' \succeq x$  for all  $0 \leq t \leq 1$ .

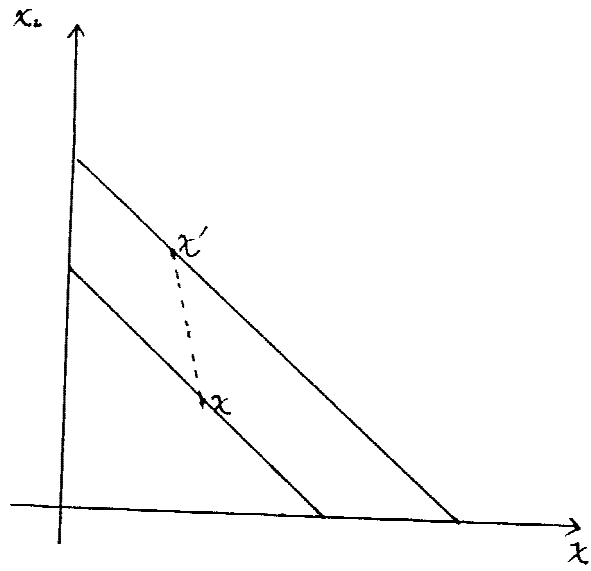


Figure 2.7: Linear indifference curves are convex, but not strict convex.

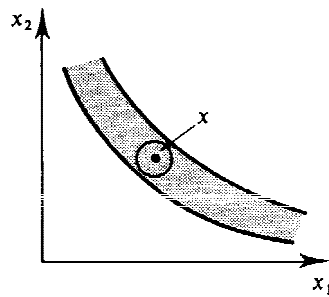


Figure 2.8: “Thick” indifference curves are weakly convex, but not convex.

**Remark 2.3.2** The convexity of preferences implies that people want to diversify their consumptions (the consumer prefers averages to extremes), and thus, convexity can be viewed as the formal expression of basic measure of economic markets for diversification. Note that convex preferences may have indifference curves that exhibit “flat spots,” while strictly convex preferences have indifference curves that are strictly rotund. The strict convexity of  $\succsim_i$  implies the neoclassical assumption of “diminishing marginal rates of substitution” between any two goods as shown in Figure 2.9.

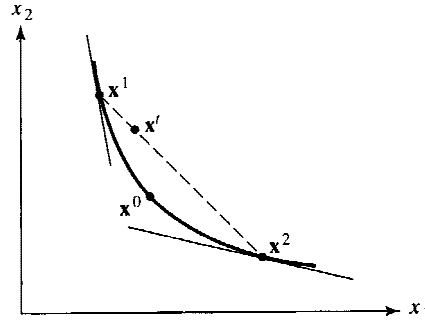


Figure 2.9: The marginal rate of substitution is diminishing when we the consumption of good 1 increases.

### 2.3.2 The Utility Function

Sometimes it is easier to work directly with the preference relation and its associated sets. But other times, especially when one wants to use calculus methods, it is easier to work with preferences that can be represented by a **utility function**; that is, a function  $u: X \rightarrow \mathbb{R}$  such that  $\mathbf{x} \succeq \mathbf{y}$  if and only if  $u(\mathbf{x}) \geq u(\mathbf{y})$ . In the following, we give some examples of utility functions.

**Example 2.3.2 (Cobb-Douglas Utility Function)** A utility function that is used frequently for illustrative and empirical purposes is the Cobb-Douglas utility function,

$$u(x_1, x_2, \dots, x_L) = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_L^{\alpha_L}$$

with  $\alpha_l > 0$ ,  $l = 1, \dots, L$ . This utility function represents a preference ordering that is continuous, strictly monotonic, and strictly convex in  $\mathbb{R}_{++}^L$ .

**Example 2.3.3 (Linear Utility Function)** A utility function that describes perfect substitution between goods is the linear utility function,

$$u(x_1, x_2, \dots, x_L) = a_1 x_1 + a_2 x_2 + \dots + a_L x_L$$

with  $a_l \geq 0$  for all  $l = 1, \dots, L$  and  $a_l > 0$  for at least  $l$ . This utility function represents a preference ordering that is continuous, monotonic, and convex in  $\mathbb{R}_+^L$ .

**Example 2.3.4 (Leontief Utility Function)** A utility function that describes perfect complement between goods is the Leontief utility function,

$$u(x_1, x_2, \dots, x_L) = \min\{a_1x_1, a_2x_2, \dots, a_Lx_L\}$$

with  $a_l \geq 0$  for all  $l = 1, \dots, L$  and  $a_l > 0$  for at least  $l$ . This represents a preference that all commodities should be used together in order to increase consumer utility. This utility function represents a preference ordering that is also continuous, monotonic, and convex in  $\mathbb{R}_+^L$ .

Not all preference orderings can be represented by utility functions, but it can be shown that any (upper semi-)continuous preference ordering can be represented by a (upper semi-)continuous utility function. We now prove a weaker version of this assertion. The following proposition shows the existence of a utility function when a preference ordering is continuous and strictly monotonic.

**Theorem 2.3.1 (Existence of a Utility Function)** *Suppose preferences are complete, reflexive, transitive, continuous, and strongly monotonic. Then there exists a continuous utility function  $u: R_+^k \rightarrow R$  which represents those preferences.*

PROOF. Let  $\mathbf{e}$  be the vector in  $R_+^k$  consisting of all ones. Then given any vector  $\mathbf{x}$  let  $u(\mathbf{x})$  be that number such that  $\mathbf{x} \sim u(\mathbf{x})\mathbf{e}$ . We have to show that such a number exists and is unique.

Let  $B = \{t \text{ in } R: t\mathbf{e} \succeq \mathbf{x}\}$  and  $W = \{t \text{ in } R: \mathbf{x} \succeq t\mathbf{e}\}$ . Then strong monotonicity implies  $B$  is nonempty;  $W$  is certainly nonempty since it contains 0. Continuity implies both sets are closed. Since the real line is connected, there is some  $t_x$  such that  $t_x\mathbf{e} \sim \mathbf{x}$ . We have to show that this utility function actually represents the underlying preferences. Let

$$u(\mathbf{x}) = t_x \quad \text{where } t_x\mathbf{e} \sim \mathbf{x}$$

$$u(\mathbf{y}) = t_y \quad \text{where } t_y\mathbf{e} \sim \mathbf{y}.$$

Then if  $t_x < t_y$ , strong monotonicity shows that  $t_x\mathbf{e} \prec t_y\mathbf{e}$ , and transitivity shows that

$$\mathbf{x} \sim t_x\mathbf{e} \prec t_y\mathbf{e} \sim \mathbf{y}.$$

Similarly, if  $\mathbf{x} \succ \mathbf{y}$ , then  $t_x\mathbf{e} \succ t_y\mathbf{e}$  so that  $t_x$  must be greater than  $t_y$ .

Finally, we show that the function  $u$  defined above is continuous. Suppose  $\{x_k\}$  is a sequence with  $x_k \rightarrow x$ . We want to show that  $u(x_k) \rightarrow u(x)$ . Suppose not. Then we can find  $\epsilon > 0$  and an infinite number of  $k$ 's such that  $u(x_k) > u(x) + \epsilon$  or an infinite set of  $k$ 's such that  $u(x_k) < u(x) - \epsilon$ . Without loss of generality, let us assume the first of these. This means that  $x_k \sim u(x_k)e \succ (u(x) + \epsilon)e \sim x + \epsilon e$ . So by transitivity,  $x_k \succ x + \epsilon e$ . But for a large  $k$  in our infinite set,  $x + \epsilon e > x_k$ , so  $x + \epsilon e \succ x_k$ , contradiction. Thus  $u$  must be continuous.

The following is an example of the non-existence of utility function when preference ordering is not continuous.

**Example 2.3.5 (Non-Representation of Lexicographic Ordering by a Function)**

Given a upper semi-continuous utility function  $u$ , the upper contour set  $\{x \in X : u(x) \geq \bar{u}\}$  must be closed for each value of  $\bar{u}$ . It follows that the lexicographic ordering defined on  $\mathbb{R}^L$  discussed earlier cannot be represented by a upper semi-continuous utility function because its upper contour sets are not closed.

The role of the utility function is to efficiently record the underlying preference ordering. The actual numerical values of  $u$  have essentially no meaning: only the sign of the difference in the value of  $u$  between two points is significant. Thus, a utility function is often a very convenient way to describe preferences, but it should not be given any psychological interpretation. The only relevant feature of a utility function is its ordinal character. Specifically, we can show that a utility function is unique only to within an arbitrary, strictly increasing transformation.

**Theorem 2.3.2 (Invariance of Utility Function to Monotonic Transforms)** *If  $u(\mathbf{x})$  represents some preferences  $\succeq$  and  $f: R \rightarrow R$  is strictly monotonic increasing, then  $f(u(\mathbf{x}))$  will represent exactly the same preferences.*

Proof. This is because  $f(u(\mathbf{x})) \geq f(u(\mathbf{y}))$  if and only if  $u(\mathbf{x}) \geq u(\mathbf{y})$ .

This invariance theorem is useful in many aspects. For instance, as it will be shown, we may use it to simplify the computation of deriving a demand function from utility maximization.

We can also use utility function to find the marginal rate of substitution between goods. Let  $u(x_1, \dots, x_k)$  be a utility function. Suppose that we increase the amount of

good  $i$ ; how does the consumer have to change his consumption of good  $j$  in order to keep utility constant?

Let  $dx_i$  and  $dx_j$  be the differentials of  $x_i$  and  $x_j$ . By assumption, the change in utility must be zero, so

$$\frac{\partial u(\mathbf{x})}{\partial x_i} dx_i + \frac{\partial u(\mathbf{x})}{\partial x_j} dx_j = 0.$$

Hence

$$\frac{dx_j}{dx_i} = -\frac{\frac{\partial u(\mathbf{x})}{\partial x_i}}{\frac{\partial u(\mathbf{x})}{\partial x_j}} \equiv -\frac{MU_{x_i}}{MU_{x_j}}$$

which gives the marginal rate of substitution between goods  $i$  and  $j$  and is defined as the ratio of the marginal utility of  $x_i$  and the marginal utility of  $x_j$ .

**Remark 2.3.3** The marginal rate of substitution does not depend on the utility function chosen to represent the underlying preferences. To prove this, let  $v(u)$  be a monotonic transformation of utility. The marginal rate of substitution for this utility function is

$$\frac{dx_j}{dx_i} = -\frac{v'(u) \frac{\partial u(\mathbf{x})}{\partial x_i}}{v'(u) \frac{\partial u(\mathbf{x})}{\partial x_j}} = -\frac{\frac{\partial u(\mathbf{x})}{\partial x_i}}{\frac{\partial u(\mathbf{x})}{\partial x_j}}.$$

The important properties of a preference ordering can be easily verified by examining utility function. The properties are summarized in the following proposition.

**Proposition 2.3.1** *Let  $\succeq$  be represented by a utility function  $u : X \rightarrow \mathbb{R}$ . Then:*

- (1) *An ordering is strictly monotonic if and only if  $u$  is strictly monotonic.*
- (2) *An ordering is continuous if and only if  $u$  is continuous.*
- (3) *An ordering is weakly convex if and only if  $u$  is quasi-concave.*
- (4) *An ordering is strictly convex if and only if  $u$  is strictly quasi-concave.*

Note that a function  $u$  is quasi-concavity if for any  $c$ ,  $u(x) \geq c$  and  $u(y) \geq c$  implies that  $u(tx + (1-t)y) \geq c$  for all  $t$  with  $0 < t < 1$ . A function  $u$  is strictly quasi-concavity if for any  $c$   $u(x) \geq c$  and  $u(y) \geq c$  implies that  $u(tx + (1-t)y) > c$  for all  $t$  with  $0 < t < 1$ .

**Remark 2.3.4** The strict quasi-concave of  $u(\mathbf{x})$  can be checked by verifying if the naturally ordered principal minors of the bordered Hessian alternate in sign, i.e.,

$$\begin{vmatrix} 0 & u_1 & u_2 \\ u_1 & u_{11} & u_{12} \\ u_2 & u_{21} & u_{22} \end{vmatrix} > 0,$$

$$\begin{vmatrix} 0 & u_1 & u_2 & u_3 \\ u_1 & u_{11} & u_{12} & u_{13} \\ u_2 & u_{21} & u_{22} & u_{23} \\ u_3 & u_{31} & u_{32} & u_{33} \end{vmatrix} < 0,$$

and so on, where  $u_i = \frac{\partial u}{\partial x_i}$  and  $u_{ij} = \frac{\partial^2 u}{\partial x_i \partial x_j}$ .

**Example 2.3.6** Suppose the preference ordering is represented by the Cobb-Douglas utility function :  $u(x_1, x_2) = x_1^\alpha x_2^\beta$  with  $\alpha > 0$  and  $\beta > 0$ . Then, we have

$$\begin{aligned} u_x &= \alpha x^{\alpha-1} y^\beta \\ u_y &= \beta x^\alpha y^{\beta-1} \\ u_{xx} &= \alpha(\alpha-1) x^{\alpha-2} y^\beta \\ u_{xy} &= \alpha\beta x^{\alpha-1} y^{\beta-1} \\ u_{yy} &= \beta(\beta-1) x^\alpha y^{\beta-2} \end{aligned}$$

and thus

$$\begin{aligned} \begin{vmatrix} 0 & u_x & u_y \\ u_x & u_{xx} & u_{xy} \\ u_y & u_{xy} & u_{yy} \end{vmatrix} &= \begin{vmatrix} 0 & \alpha x^{\alpha-1} y^\beta & \beta x^\alpha y^{\beta-1} \\ \alpha x^{\alpha-1} y^\beta & \alpha(\alpha-1) x^{\alpha-2} y^\beta & \alpha\beta x^{\alpha-1} y^{\beta-1} \\ \beta x^\alpha y^{\beta-1} & \alpha\beta x^{\alpha-1} y^{\beta-1} & \beta(\beta-1) x^\alpha y^{\beta-2} \end{vmatrix} \\ &= x^{3\alpha-2} y^{3\beta-2} [\alpha\beta(\alpha+\beta)] > 0 \quad \text{for all } (x, y) > 0, \end{aligned}$$

which means  $u$  is strictly quasi-concave in  $\mathbb{R}_{++}^2$ .



## 2.4 Utility Maximization and Optimal Choice

### 2.4.1 Consumer Behavior: Utility Maximization

A foundational hypothesis on individual behavior in modern economics in general and the consumer theory in particular is that a rational agent will always choose a most preferred bundle from the set of affordable alternatives. We will derive demand functions by considering a model of utility-maximizing behavior coupled with a description of underlying economic constraints.

### 2.4.2 Consumer's Optimal Choice

In the basic problem of preference maximization, the set of affordable alternatives is just the set of all bundles that satisfy the consumer's budget constraint we discussed before. That is, the problem of preference maximization can be written as:

$$\begin{aligned} & \max u(\mathbf{x}) \\ \text{such that} & \quad \mathbf{p}\mathbf{x} \leq m \\ & \mathbf{x} \text{ is in } X. \end{aligned}$$

There will exist a solution to this problem if the utility function is continuous and that the constraint set is closed and bounded. The constraint set is certainly closed. If  $p_i > 0$  for  $i = 1, \dots, k$  and  $m > 0$ , it is not difficult to show that the constraint set will be bounded. If some price is zero, the consumer might want an infinite amount of the corresponding good.

**Proposition 2.4.1** *Under the local nonsatiation assumption, a utility-maximizing bundle  $\mathbf{x}^*$  must meet the budget constraint with equality.*

Proof. Suppose we get an  $\mathbf{x}^*$  where  $\mathbf{p}\mathbf{x}^* < m$ . Since  $\mathbf{x}^*$  costs strictly less than  $m$ , every bundle in  $X$  close enough to  $\mathbf{x}^*$  also costs less than  $m$  and is therefore feasible. But, according to the local nonsatiation hypothesis, there must be some bundle  $\mathbf{x}$  which is close to  $\mathbf{x}^*$  and which is preferred to  $\mathbf{x}^*$ . But this means that  $\mathbf{x}^*$  could not maximize preferences on the budget set  $B(p, m)$ .

This proposition allows us to restate the consumer's problem as

$$\begin{aligned} & \max u(\mathbf{x}) \\ & \text{such that } \mathbf{p}\mathbf{x} = m. \end{aligned}$$

The value of  $\mathbf{x}$  that solves this problem is the consumer's **demanded bundle**: it expresses how much of each good the consumer desires at a given level of prices and income. In general, the optimal consumption is not unique. Denote by  $\mathbf{x}(\mathbf{p}, m)$  the set of all utility maximizing consumption bundles and it is called the consumer's **demand correspondence**. When there is a unique demanded bundle for each  $(p, m)$ ,  $\mathbf{x}(\mathbf{p}, m)$  becomes a function and thus is called the consumer's **demand function**. We will see from the following proposition that strict convexity of preferences will ensure the uniqueness of optimal bundle.

**Proposition 2.4.2 (Uniqueness of Demanded Bundle)** *If preferences are strictly convex, then for each  $p > 0$  there is a unique bundle  $x$  that maximizes  $u$  on the consumer's budget set,  $B(p, m)$ .*

PROOF. Suppose  $\mathbf{x}'$  and  $\mathbf{x}''$  both maximize  $u$  on  $B(\mathbf{p}, m)$ . Then  $\frac{1}{2}\mathbf{x}' + \frac{1}{2}\mathbf{x}''$  is also in  $B(\mathbf{p}, m)$  and is strictly preferred to  $\mathbf{x}'$  and  $\mathbf{x}''$ , which is a contradiction.

Since multiplying all prices and income by some positive number does not change the budget set at all and thus cannot change the answer to the utility maximization problem.

**Proposition 2.4.3 (Homogeneity of Demand Function)** *The consumer's demand function  $x(p, m)$  is homogeneous of degree 0 in  $(\mathbf{p}, m) > 0$ , i.e.,  $x(tp, tm) = x(p, m)$ .*

Note that a function  $f(\mathbf{x})$  is **homogeneous of degree  $k$**  if  $f(t\mathbf{x}) = t^k f(\mathbf{x})$  for all  $t > 0$ .

### 2.4.3 Consumer's First Order-Conditions

We can characterize optimizing behavior by calculus, as long as the utility function is differentiable. We will analyze this constrained maximization problem using the method of Lagrange multipliers. The Lagrangian for the utility maximization problem can be written as

$$\mathcal{L} = u(\mathbf{x}) - \lambda(\mathbf{p}\mathbf{x} - m),$$

where  $\lambda$  is the Lagrange multiplier. Suppose preference is locally non-satiated. Differentiating the Lagrangian with respect to  $x_i$ , gives us the first-order conditions for the interior solution

$$\frac{\partial u(\mathbf{x})}{\partial x_i} - \lambda p_i = 0 \quad \text{for } i = 1, \dots, L \quad (2.1)$$

$$px = m \quad (2.2)$$

Using vector notation, we can also write equation (2.1) as

$$\mathbf{D}u(\mathbf{x}) = \lambda \mathbf{p}.$$

Here

$$\mathbf{D}u(x) = \left( \frac{\partial u(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial u(\mathbf{x})}{\partial x_L} \right)$$

is the **gradient** of  $u$ : the vector of partial derivatives of  $u$  with respect to each of its arguments.

In order to interpret these conditions we can divide the  $i^{th}$  first-order condition by the  $j^{th}$  first-order condition to eliminate the Lagrange multiplier. This gives us

$$\frac{\frac{\partial u(\mathbf{x}^*)}{\partial x_i}}{\frac{\partial u(\mathbf{x}^*)}{\partial x_j}} = \frac{p_i}{p_j} \quad \text{for } i, j = 1, \dots, L. \quad (2.3)$$

The fraction on the left is the marginal rate of substitution between good  $i$  and  $j$ , and the fraction on the right is economic rate of substitution between goods  $i$  and  $j$ . Maximization implies that these two rates of substitution should be equal. Suppose they were not; for example, suppose

$$\frac{\frac{\partial u(\mathbf{x}^*)}{\partial x_i}}{\frac{\partial u(\mathbf{x}^*)}{\partial x_j}} = \frac{1}{1} \neq \frac{2}{1} = \frac{p_i}{p_j}. \quad (2.4)$$

Then, if the consumer gives up one unit of good  $i$  and purchases one unit of good  $j$ , he or she will remain on the same indifference curve and have an extra dollar to spend. Hence, total utility can be increased, contradicting maximization.

Figure 2.10 illustrates the argument geometrically. The budget line of the consumer is given by  $\{\mathbf{x}: p_1x_1 + p_2x_2 = m\}$ . This can also be written as the graph of an implicit function:  $x_2 = m/p_2 - (p_1/p_2)x_1$ . Hence, the budget line has slope  $-p_1/p_2$  and vertical intercept  $m/p_2$ . The consumer wants to find the point on this budget line that achieves

highest utility. This must clearly satisfy the tangency condition that the slope of the indifference curve equals the slope of the budget line so that the marginal rate of substitution of  $x_1$  for  $x_2$  equals the economic rate of substitution of  $x_1$  for  $x_2$ .

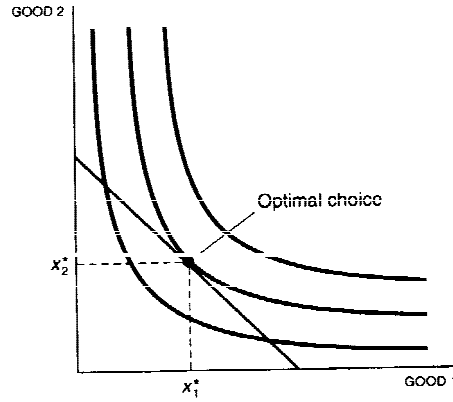


Figure 2.10: Preference maximization. The optimal consumption bundle will be at a point where an indifference curve is tangent to the budget constraint.

**Remark 2.4.1** The calculus conditions derived above make sense only when the choice variables can be varied in an open neighborhood of the optimal choice and the budget constraint is binding. In many economic problems the variables are naturally nonnegative. If some variables have a value of zero at the optimal choice, the calculus conditions described above may be inappropriate. The necessary modifications of the conditions to handle **boundary solutions** are not difficult to state. The relevant first-order conditions are given by means of the so-called Kuhn-Tucker conditions:

$$\frac{\partial u(\mathbf{x})}{\partial x_i} - \lambda p_i \leq 0 \quad \text{with equality if } x_i > 0 \quad i = 1, \dots, L. \quad (2.5)$$

$$px \leq m \quad \text{with equality if } \lambda > 0 \quad (2.6)$$

Thus the marginal utility from increasing  $x_i$  must be less than or equal to  $\lambda p_i$ , otherwise the consumer would increase  $x_i$ . If  $x_i = 0$ , the marginal utility from increasing  $x_i$  may be less than  $\lambda p_i$  — which is to say, the consumer would like to decrease  $x_i$ . But since  $x_i$  is already zero, this is impossible. Finally, if  $x_i > 0$  so that the nonnegativity constraint is not binding, we will have the usual conditions for an interior solution.

## 2.4.4 Sufficiency of Consumer's First-Order Conditions

The above first-order conditions are merely necessary conditions for a local optimum. However, for the particular problem at hand, these necessary first-order conditions are in fact sufficient for a global optimum when a utility function is quasi-concave. We then have the following proposition.

**Proposition 2.4.4** *Suppose that  $u(x)$  is differentiable and quasi-concave on  $\mathbb{R}_+^L$  and  $(p, m) > 0$ . If  $(x, \lambda)$  satisfies the first-order conditions given in (2.5) and (2.6), then  $x$  solves the consumer's utility maximization problem at prices  $p$  and income  $m$ .*

PROOF. Since the budget set  $B(p, m)$  is convex and  $u(x)$  is differentiable and quasi-concave on  $\mathbb{R}_+^L$ , by Proposition 1.2.1, we know  $x$  solves the consumer's utility maximization problem at prices  $p$  and income  $m$ .

With the sufficient conditions in hand, it is enough to find a solution  $(x, \lambda)$  that satisfies the first-order conditions (2.5) and (2.6). The conditions can typically be used to solve the demand functions  $x_i(p, m)$  as we show in the following examples.

**Example 2.4.1** Suppose the preference ordering is represented by the Cobb-Douglas utility function:  $u(x_1, x_2) = x_1^a x_2^{1-a}$ , which is strictly quasi-concave on  $\mathbb{R}_{++}^2$ . Since any monotonic transform of this function represents the same preferences, we can also write  $u(x_1, x_2) = a \ln x_1 + (1 - a) \ln x_2$ .

The demand functions can be derived by solving the following problem:

$$\begin{aligned} \max \quad & a \ln x_1 + (1 - a) \ln x_2 \\ \text{such that} \quad & p_1 x_1 + p_2 x_2 = m. \end{aligned}$$

The first-order conditions are

$$\frac{a}{x_1} - \lambda p_1 = 0$$

or

$$\frac{a}{p_1 x_1} = \frac{1 - a}{p_2 x_2}.$$

Cross multiply and use the budget constraint to get

$$\begin{aligned} ap_2 x_2 &= p_1 x_1 - ap_1 x_1 \\ am &= p_1 x_1 \\ x_1(p_1, p_2, m) &= \frac{am}{p_1}. \end{aligned}$$

Substitute into the budget constraint to get the demand function for the second commodity:

$$x_2(p_1, p_2, m) = \frac{(1-a)m}{p_2}.$$

**Example 2.4.2** Suppose the preference ordering is represented by the Leontief utility function :  $u(x_1, x_2) = \min\{ax_1, bx_2\}$ . Since the Leontief utility function is not differentiable, so the maximum must be found by a direct argument. Assume  $p > 0$ .

The optimal solution must be at the kink point of the indifference curve. That is,

$$ax_1 = bx_2.$$

Substituting  $x_1 = \frac{b}{a}x_2$  into the budget constraint  $px = m$ , we have

$$p_1 \frac{b}{a} x_2 + p_2 x_2 = m$$

and thus the demand functions are given by

$$x_2(p_1, p_2, m) = \frac{am}{bp_1 + ap_2}$$

and

$$x_1(p_1, p_2, m) = \frac{bm}{bp_1 + ap_2}$$

**Example 2.4.3** Now suppose the preference ordering is represented by the linear utility function :

$$u(x, y) = ax + by.$$

Since the marginal rate of substitution is  $a/b$  and the economic rate of substitution is  $p_x/p_y$  are both constant, they cannot be in general equal. So the first-order condition cannot hold with equality as long as  $a/b \neq p_x/p_y$ . In this case the answer to the utility-maximization problem typically involves a boundary solution: only one of the two goods will be consumed. It is worthwhile presenting a more formal solution since it serves as a nice example of the Kuhn-Tucker theorem in action. The Kuhn-Tucker theorem is the appropriate tool to use here, since we will almost never have an interior solution.

The Lagrange function is

$$L(x, y, \lambda) = ax + by + \lambda(m - p_x x - p_y y)$$

and thus

$$\frac{\partial L}{\partial x} = a - \lambda p_x \quad (2.7)$$

$$\frac{\partial L}{\partial y} = b - \lambda p_y \quad (2.8)$$

$$\frac{\partial L}{\partial \lambda} = m - p_x x - p_y y \quad (2.9)$$

$$(2.10)$$

There are four cases to be considered:

Case 1.  $x > 0$  and  $y > 0$ . Then we have  $\frac{\partial L}{\partial x} = 0$  and  $\frac{\partial L}{\partial y} = 0$ . Thus,  $\frac{a}{b} = \frac{p_x}{p_y}$ . Since  $\lambda = \frac{a}{p_x} > 0$ , we have  $p_x x + p_y y = m$  and thus all  $x$  and  $y$  that satisfy  $p_x x + p_y y = m$  are the optimal consumptions.

Case 2.  $x > 0$  and  $y = 0$ . Then we have  $\frac{\partial L}{\partial x} = 0$  and  $\frac{\partial L}{\partial y} \leq 0$ . Thus,  $\frac{a}{b} \geq \frac{p_x}{p_y}$ . Since  $\lambda = \frac{a}{p_x} > 0$ , we have  $p_x x + p_y y = m$  and thus  $x = \frac{m}{p_x}$  is the optimal consumption.

Case 3.  $x = 0$  and  $y > 0$ . Then we have  $\frac{\partial L}{\partial x} \leq 0$  and  $\frac{\partial L}{\partial y} = 0$ . Thus,  $\frac{a}{b} \leq \frac{p_x}{p_y}$ . Since  $\lambda = \frac{b}{p_y} > 0$ , we have  $p_x x + p_y y = m$  and thus  $y = \frac{m}{p_y}$  is the optimal consumption.

Case 4.  $x = 0$  and  $y = 0$ . Then we have  $\frac{\partial L}{\partial x} \leq 0$  and  $\frac{\partial L}{\partial y} \leq 0$ . Since  $\lambda \geq \frac{b}{p_y} > 0$ , we have  $p_x x + p_y y = m$  and thus  $m = 0$  because  $x = 0$  and  $y = 0$ .

In summary, the demand functions are given by

$$(x(p_x, p_y, m), y(p_x, p_y, m)) = \begin{cases} (m/p_x, 0) & \text{if } a/b > p_x/p_y \\ (0, m/p_y) & \text{if } a/b < p_x/p_y \\ (x, m/p_x - p_y/p_x x) & \text{if } a/b = p_x/p_y \end{cases}$$

for all  $x \in [0, m/p_x]$ .

**Remark 2.4.2** In fact, it is easily found out the optimal solutions by comparing relatives steepness of the indifference curves and the budget line. For instance, as shown in Figure 2.11 below, when  $a/b > p_x/p_y$ , the indifference curves become steeper, and thus the optimal solution is the one the consumer spends his all income on good  $x$ . When  $a/b < p_x/p_y$ , the indifference curves become flatter, and thus the optimal solution is the one the consumer spends his all income on good  $y$ . When  $a/b = p_x/p_y$ , the indifference curves and the budget line are parallel and coincide at the optimal solutions, and thus the optimal solutions are given by all the points on the budget line.

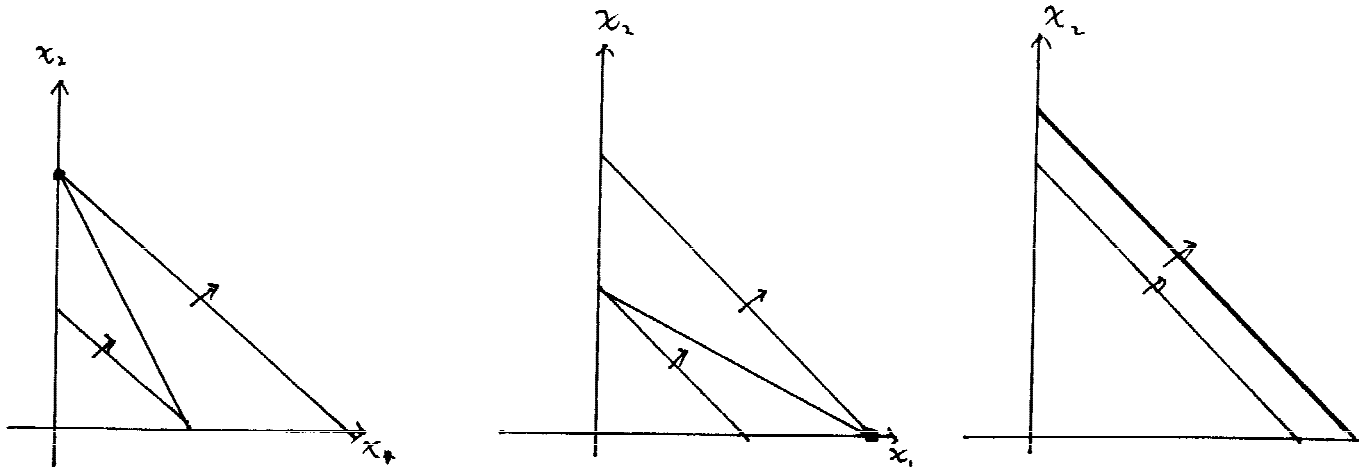


Figure 2.11: Maximize utility and minimize expenditure are normally equivalent.

## 2.5 Indirect Utility, and Expenditure, and Money Metric Utility Functions

### 2.5.1 The Indirect Utility Function

The ordinary utility function,  $u(x)$ , is defined over the consumption set  $X$  and therefore to as the **direct utility function**. Given prices  $p$  and income  $m$ , the consumer chooses a utility-maximizing bundle  $x(p, m)$ . The level of utility achieved when  $x(p, m)$  is chosen thus will be the highest level permitted by the consumer's budget constraint facing  $p$  and  $m$ , and can be denoted by

$$v(\mathbf{p}, m) = \max u(\mathbf{x})$$

such that  $\mathbf{p}\mathbf{x} = m$ .

The function  $v(\mathbf{p}, m)$  that gives us the maximum utility achievable at given prices and income is called the **indirect utility function** and thus it is a compose function of  $u(\cdot)$  and  $x(p, m)$ , i.e.,

$$v(p, m) = u(x(p, m)). \quad (2.11)$$

The properties of the indirect utility function are summarized in the following proposition.



**Proposition 2.5.1 (Properties of the indirect utility function)** *If  $u(x)$  is continuous and monotonic on  $\mathbb{R}_+^L$  and  $(p, m) > 0$ , the indirect utility function has the following properties:*

- (1)  $v(\mathbf{p}, m)$  is nonincreasing in  $\mathbf{p}$ ; that is, if  $\mathbf{p}' \geq \mathbf{p}$ ,  $v(\mathbf{p}', m) \leq v(\mathbf{p}, m)$ . Similarly,  $v(\mathbf{p}, m)$  is nondecreasing in  $m$ .
- (2)  $v(\mathbf{p}, m)$  is homogeneous of degree 0 in  $(\mathbf{p}, m)$ .
- (3)  $v(\mathbf{p}, m)$  is quasiconvex in  $\mathbf{p}$ ; that is,  $\{\mathbf{p}: v(\mathbf{p}, m) \leq k\}$  is a convex set for all  $k$ .
- (4)  $v(\mathbf{p}, m)$  is continuous at all  $\mathbf{p} \gg 0, m > 0$ .

PROOF.

- (1) Let  $B = \{\mathbf{x}: \mathbf{p}\mathbf{x} \leq m\}$  and  $B' = \{\mathbf{x}: \mathbf{p}'\mathbf{x} \leq m\}$  for  $\mathbf{p}' \geq \mathbf{p}$ . Then  $B'$  is contained in  $B$ . Hence, the maximum of  $u(\mathbf{x})$  over  $B$  is at least as big as the maximum of  $u(\mathbf{x})$  over  $B'$ . The argument for  $m$  is similar.
- (2) If prices and income are both multiplied by a positive number, the budget set doesn't change at all. Thus,  $v(t\mathbf{p}, tm) = v(\mathbf{p}, m)$  for  $t > 0$ .
- (3) Suppose  $\mathbf{p}$  and  $\mathbf{p}'$  are such that  $v(\mathbf{p}, m) \leq k, v(\mathbf{p}', m) \leq k$ . Let  $\mathbf{p}'' = t\mathbf{p} + (1 - t)\mathbf{p}'$ . We want to show that  $v(\mathbf{p}'', m) \leq k$ . Define the budget sets:

$$\begin{aligned} B &= \{\mathbf{x}: \mathbf{p}\mathbf{x} \leq m\} \\ B' &= \{\mathbf{x}: \mathbf{p}'\mathbf{x} \leq m\} \\ B'' &= \{\mathbf{x}: \mathbf{p}''\mathbf{x} \leq m\} \end{aligned}$$

We will show that any  $\mathbf{x}$  in  $B''$  must be in either  $B$  or  $B'$ ; that is, that  $B \cup B' \supset B''$ . Suppose not. We then must have  $\mathbf{p}\mathbf{x} > m$  and  $\mathbf{p}'\mathbf{x} > m$ . Multiplying the first inequality by  $t$  and the second by  $(1 - t)$  and then summing, we find that  $t\mathbf{p}\mathbf{x} + (1 - t)\mathbf{p}'\mathbf{x} > m$  which contradicts our original assumption. Then  $B''$  is contained in  $B \cup B'$ , the maximum of  $u(\mathbf{x})$  over  $B''$  is at most as big as the maximum of  $u(\mathbf{x})$  over  $B \cup B'$ , and thus  $v(\mathbf{p}'', m) \leq k$  by noting that  $v(\mathbf{p}, m) \leq k$  and  $v(\mathbf{p}', m) \leq k$ .

(4) This follows from the Maximum Theorem.

**Example 2.5.1 (The General Cobb-Douglas Utility Function)** Suppose a preference ordering is represented by the Cobb-Douglas utility function is given by:

$$u(x) = \prod_{l=1}^L (x_l)^{\alpha_l}, \quad \alpha_l > 0, l = 1, 2, \dots, L.$$

Since any monotonic transform of this function represents the same preferences, we can also write

$$u(x) = \prod_{l=1}^L (x_l)^{\frac{\alpha_l}{\alpha}}, \quad \alpha_l > 0, l = 1, 2, \dots, L.$$

where  $\alpha = \sum_{l=1}^L \alpha_l$ . Let  $a_l = \alpha_l / \alpha$ . Then it reduces to the Cobb-Douglas utility we examined before and thus the demand functions are given by

$$x_l(p, m) = \frac{a_l m}{p_l} = \frac{\alpha_l m}{\alpha p_l}, \quad l = 1, 2, \dots, L.$$

Substitute into the objective function and eliminate constants to get the indirect utility function:

$$v(p, m) = \prod_{l=1}^L \left( \frac{\alpha_l m}{\alpha p_l} \right)^{\alpha_l}$$

The above example also shows that monotonic transformation sometimes is very useful to simplify the computation of finding solutions.

## 2.5.2 The Expenditure Function and Hicksian Demand

We note that if preferences satisfy the local nonsatiation assumption, then  $v(\mathbf{p}, m)$  will be *strictly* increasing in  $m$ . We then can invert the function and solve for  $m$  as a function of the level of utility; that is, given any level of utility,  $u$ , we can find the minimal amount of income necessary to achieve utility  $u$  at prices  $\mathbf{p}$ . The function that relates income and utility in this way—the inverse of the indirect utility function—is known as the **expenditure function** and is denoted by  $e(\mathbf{p}, u)$ . Formally, the expenditure function is given by the following problem:

$$\begin{aligned} e(\mathbf{p}, u) &= \min \mathbf{p}\mathbf{x} \\ \text{such that } u(\mathbf{x}) &\geq u. \end{aligned}$$

The expenditure function gives the minimum cost of achieving a fixed level of utility. The solution which is the function of  $(p, u)$  is denoted by  $\mathbf{h}(\mathbf{p}, u)$  and called the **Hicksian demand function**. The Hicksian demand function tells us what consumption bundle achieves a target level of utility and minimizes total expenditure.

A Hicksian demand function is sometimes called a compensated demand function. This terminology comes from viewing the demand function as being constructed by varying prices and *income* so as to keep the consumer at a fixed level of utility. Thus, the income changes are arranged to “compensate” for the price changes. Hicksian demand functions are not directly observable since they depend on utility, which is not directly observable. Demand functions expressed as a function of prices and income are observable; when we want to emphasize the difference between the Hicksian demand function and the usual demand function, we will refer to the latter as the **Marshallian demand function**,  $\mathbf{x}(\mathbf{p}, m)$ . The Marshallian demand function is just the ordinary market demand function we have been discussing all along.

**Proposition 2.5.2** [*Properties of the Expenditure Function.*] *If  $u(x)$  is continuous and locally non-satiated on  $\mathbb{R}_+^L$  and  $(p, m) > 0$ , the expenditure function has the following properties:*

- (1)  $e(\mathbf{p}, u)$  is nondecreasing in  $\mathbf{p}$ .
- (2)  $e(\mathbf{p}, u)$  is homogeneous of degree 1 in  $\mathbf{p}$ .
- (3)  $e(\mathbf{p}, u)$  is concave in  $\mathbf{p}$ .
- (4)  $e(\mathbf{p}, u)$  is continuous in  $\mathbf{p}$ , for  $p \gg 0$ .
- (5) For all  $p > 0$ ,  $e(p, \mathbf{u})$  is strictly increasing in  $\mathbf{u}$ .
- (6) *Shephard's lemma: If  $\mathbf{h}(\mathbf{p}, u)$  is the expenditure-minimizing bundle necessary to achieve utility level  $u$  at prices  $\mathbf{p}$ , then  $h_i(\mathbf{p}, u) = \frac{\partial e(\mathbf{p}, u)}{\partial p_i}$  for  $i = 1, \dots, k$  assuming the derivative exists and that  $p_i > 0$ .*

**PROOF.** Since the expenditure function is the inverse function of the indirect utility function, Properties (1), (4)-(5) are true by Properties (1) and (4) of the indirect utility given in Proposition 2.5.1. We only need to show the Properties (2), (3) and (6).

- (2) We show that if  $\mathbf{x}$  is the expenditure-minimizing bundle at prices  $p$ , then  $x$  also minimizes the expenditure at prices  $tp$ . Suppose not, and let  $x'$  be an expenditure minimizing bundle at  $tp$  so that  $tpx' < tp\mathbf{x}$ . But this inequality implies  $\mathbf{p}x' < \mathbf{p}\mathbf{x}$ , which contradicts the definition of  $\mathbf{x}$ . Hence, multiplying prices by a positive scalar  $t$  does not change the composition of an expenditure minimizing bundle, and, thus, expenditures must rise by exactly a factor of  $t$ :  $e(p, u) = t\mathbf{p}\mathbf{x} = te(p, u)$ .
- (3) Let  $(\mathbf{p}, \mathbf{x})$  and  $(\mathbf{p}', \mathbf{x}')$  be two expenditure-minimizing price-consumption combinations and let  $\mathbf{p}'' = t\mathbf{p} + (1 - t)\mathbf{p}'$  for any  $0 \leq t \leq 1$ . Now,

$$e(\mathbf{p}'', u) = \mathbf{p}''\mathbf{x}'' = t\mathbf{p}\mathbf{x}'' + (1 - t)\mathbf{p}'\mathbf{x}''.$$

Since  $\mathbf{x}''$  is not necessarily the minimal expenditure to reach  $u$  at prices  $\mathbf{p}'$  or  $\mathbf{p}$ , we have  $\mathbf{p}\mathbf{x}'' \geq e(\mathbf{p}, u)$  and  $\mathbf{p}' \cdot \mathbf{x}'' \geq e(\mathbf{p}', u)$ . Thus,

$$e(\mathbf{p}'', u) \geq te(\mathbf{p}, u) + (1 - t)e(\mathbf{p}', u).$$

- (6) Let  $\mathbf{x}^*$  be an expenditure-minimizing bundle to achieve utility level  $u$  at prices  $\mathbf{p}^*$ . Then define the function

$$g(\mathbf{p}) = e(\mathbf{p}, u) - \mathbf{p}\mathbf{x}^*.$$

Since  $e(\mathbf{p}, u)$  is the cheapest way to achieve  $u$ , this function is always non-positive. At  $\mathbf{p} = \mathbf{p}^*$ ,  $g(\mathbf{p}^*) = 0$ . Since this is a maximum value of  $g(\mathbf{p})$ , its derivative must vanish:

$$\frac{\partial g(\mathbf{p}^*)}{\partial p_i} = \frac{\partial e(\mathbf{p}^*, u)}{\partial p_i} - x_i^* = 0 \quad i = 1, \dots, L.$$

Hence, the expenditure-minimizing bundles are just given by the vector of derivatives of the expenditure function with respect to the prices.

**Remark 2.5.1** We can also prove property (6) by applying the Envelop Theorem for the constrained version. In this problem the parameter  $a$  can be chosen to be one of the prices,  $p_i$ . Define the Lagrange function  $L(x, \lambda) = px - \lambda(u - u(x))$ . The optimal value function is the expenditure function  $e(\mathbf{p}, u)$ . The envelope theorem asserts that

$$\frac{\partial e(\mathbf{p}, u)}{\partial p_i} = \frac{\partial L}{\partial p_i} = x_i \bigg|_{x_i = h_i(\mathbf{p}, u)} = h_i(\mathbf{p}, u),$$

which is simply Shephard's lemma.

We now give some basic properties of Hicksian demand functions:

**Proposition 2.5.3 (Negative Semi-Definite Substitution Matrix)** *The matrix of substitution terms  $(\partial h_j(\mathbf{p}, u)/\partial p_i)$  is negative semi-definite.*

Proof. This follows

$$\partial h_j(p, u)/\partial p_i = \partial^2 e(\mathbf{p}, u)/\partial p_i \partial p_j,$$

which is negative semi-definite because the expenditure function is concave.

Since the substitution matrix is negative semi-definite, thus it is symmetric and has non-positive diagonal terms. Then we have

**Proposition 2.5.4 (Symmetric Substitution Terms)** *The matrix of substitution terms is symmetric, i.e.,*

$$\frac{\partial h_j(p, u)}{\partial p_i} = \frac{\partial^2 e(\mathbf{p}, u)}{\partial p_j \partial p_i} = \frac{\partial^2 e(\mathbf{p}, u)}{\partial p_i \partial p_j} = \frac{\partial h_i(\mathbf{p}, u)}{\partial p_j}.$$

**Proposition 2.5.5 (Negative Own-Substitution Terms)** *The compensated own-price effect is non-positive; that is, the Hicksian demand curves slope downward:*

$$\frac{\partial h_i(\mathbf{p}, u)}{\partial p_i} = \frac{\partial^2 e(\mathbf{p}, u)}{\partial p_i^2} \leq 0,$$

### 2.5.3 The Money Metric Utility Functions

There is a nice construction involving the expenditure function that comes up in a variety of places in welfare economics. Consider some prices  $\mathbf{p}$  and some given bundle of goods  $\mathbf{x}$ . We can ask the following question: how much money would a given consumer need at the prices  $\mathbf{p}$  to be as well off as he could be by consuming the bundle of goods  $\mathbf{x}$ ? If we know the consumer's preferences, we can simply solve the following problem:

$$\begin{aligned} m(p, x) &\equiv \min_z \mathbf{p}z \\ \text{such that } u(\mathbf{z}) &\geq u(\mathbf{x}). \end{aligned}$$

That is,

$$m(\mathbf{p}, \mathbf{x}) \equiv e(\mathbf{p}, u(\mathbf{x})).$$

This type of function is called **money metric utility function**. It is also known as the “minimum income function,” the “direct compensation function,” and by a variety

of other names. Since, for fixed  $\mathbf{p}$ ,  $m(\mathbf{p}, \mathbf{x})$  is simply a monotonic transform of the utility function and is itself a utility function.

There is a similar construct for indirect utility known as the **money metric indirect utility function**, which is given by

$$\mu(\mathbf{p}; \mathbf{q}, m) \equiv e(p, \nu(q, m)).$$

That is,  $\mu(\mathbf{p}; \mathbf{q}, m)$  measures how much money one would need at prices  $\mathbf{p}$  to be as well off as one would be facing prices  $\mathbf{q}$  and having income  $m$ . Just as in the direct case,  $\mu(\mathbf{p}; \mathbf{q}, m)$  is simply a monotonic transformation of an indirect utility function.

**Example 2.5.2 (The CES Utility Function)** The CES utility function is given by  $u(x_1, x_2) = (x_1^\rho + x_2^\rho)^{1/\rho}$ , where  $0 \neq \rho < 1$ . It can be easily verified this utility function is strictly monotonic increasing and strictly concave. Since preferences are invariant with respect to monotonic transforms of utility, we could just as well choose  $u(x_1, x_2) = \frac{1}{\rho} \ln(x_1^\rho + x_2^\rho)$ .

The first-order conditions are

$$\begin{aligned} \frac{x_1^{\rho-1}}{x_1^\rho + x_2^\rho} - \lambda p_1 &= 0 \\ \frac{x_2^{\rho-1}}{x_1^\rho + x_2^\rho} - \lambda p_2 &= 0 \\ p_1 x_1 + p_2 x_2 &= m \end{aligned}$$

Dividing the first equation by the second equation and then solving for  $x_2$ , we have

$$x_2 = x_1 \left( \frac{p_2}{p_1} \right)^{\frac{1}{\rho-1}}.$$

Substituting the above equation in the budget line and solving for  $x_1$ , we obtain

$$x_1(p, m) = \frac{p_1^{\frac{1}{\rho-1}} m}{p_1^{\frac{\rho}{\rho-1}} + p_2^{\frac{\rho}{\rho-1}}}$$

and thus

$$x_2(p, m) = \frac{p_2^{\frac{1}{\rho-1}} m}{p_1^{\frac{\rho}{\rho-1}} + p_2^{\frac{\rho}{\rho-1}}}$$

Substituting the demand functions into the utility function, we get the indirect CES utility function:

$$v(\mathbf{p}, m) = (p_1^{\rho/(\rho-1)} + p_2^{\rho/(\rho-1)})^{1-\rho/\rho} m$$

or

$$v(\mathbf{p}, m) = (p_1^r + p_2^r)^{-1/r} m$$

where  $r = \rho/(\rho - 1)$ . Inverting the above equation, we obtain the expenditure function for the CES utility function which has the form

$$e(\mathbf{p}, u) = (p_1^r + p_2^r)^{1/r} u.$$

Consequently, the money metric direct and indirect utility functions are given by

$$m(\mathbf{p}, x) = (p_1^r + p_2^r)^{1/r} (x_1^\rho + x_2^\rho)^{1/\rho}$$

and

$$\mu(\mathbf{p}; \mathbf{q}, m) = (p_1^r + p_2^r)^{1/r} (q_1^r + q_2^r)^{-1/r} m.$$

**Remark 2.5.2** The CES utility function contains several other well-known utility functions as special cases, depending on the value of the parameter  $\rho$ .

- (1) The linear utility function ( $\rho = 1$ ). Simple substitution yields

$$y = x_1 + x_2.$$

- (2) The Cobb-Douglas utility function ( $\rho = 0$ ). When  $\rho = 0$  the CES utility function is not defined, due to division by zero. However, we will show that as  $\rho$  approaches zero, the indifference curves of the CES utility function look very much like the indifference curves of the Cobb-Douglas utility function.

This is easiest to see using the marginal rate of substitution. By direct calculation,

$$MRS = - \left( \frac{x_1}{x_2} \right)^{\rho-1}. \quad (2.12)$$

As  $\rho$  approaches zero, this tends to a limit of

$$MRS = - \frac{x_2}{x_1},$$

which is simply the  $MRS$  for the Cobb-Douglas utility function.

(3) The Leontief utility function ( $\rho = -\infty$ ). We have just seen that the *MRS* of the CES utility function is given by equation (2.12). As  $\rho$  approaches  $-\infty$ , this expression approaches

$$MRS = - \left( \frac{x_1}{x_2} \right)^{-\infty} = - \left( \frac{x_2}{x_1} \right)^{\infty}.$$

If  $x_2 > x_1$  the *MRS* is (negative) infinity; if  $x_2 < x_1$  the *MRS* is zero. This means that as  $\rho$  approaches  $-\infty$ , a CES indifference curves looks like an indifference curves associated with the Leontief utility function.

#### 2.5.4 Some Important Identities

There are some important identities that tie together the expenditure function, the indirect utility function, the Marshallian demand function, and the Hicksian demand function.

Let us consider the utility maximization problem

$$\begin{aligned} v(\mathbf{p}, m^*) &= \max_{\mathbf{x}} u(\mathbf{x}) \\ \text{such that } \mathbf{p}\mathbf{x} &\leq m^*. \end{aligned} \tag{2.13}$$

Let  $\mathbf{x}^*$  be the solution to this problem and let  $u^* = u(\mathbf{x}^*)$ . Consider the expenditure minimization problem

$$\begin{aligned} e(\mathbf{p}, u^*) &= \min_{\mathbf{x}} \mathbf{p}\mathbf{x} \\ \text{such that } u(\mathbf{x}) &\geq u^*. \end{aligned} \tag{2.14}$$

An inspection of Figure 2.12 should convince you that the answers to these two problems should be the same  $\mathbf{x}^*$ . Formally, we have the following proposition.

**Proposition 2.5.6 (Equivalence of Utility Max and Expenditure Min)** *Suppose the utility function  $u$  is continuous and locally non-satiated, and suppose that  $m > 0$ . If the solutions both problems exist, then the above two problems have the same solution  $\mathbf{x}^*$ . That is,*

- 1 **Utility maximization implies expenditure minimization:** Let  $\mathbf{x}^*$  be a solution to (2.13), and let  $u = u(\mathbf{x}^*)$ . Then  $\mathbf{x}^*$  solves (2.14).



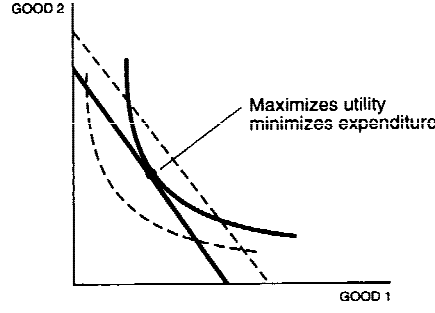


Figure 2.12: Maximize utility and minimize expenditure are normally equivalent.

**2 Expenditure minimization implies utility maximization.** *Suppose that the above assumptions are satisfied and that  $\mathbf{x}^*$  solves (2.14). Let  $m = \mathbf{p}\mathbf{x}^*$ . Then  $\mathbf{x}^*$  solves (2.13).*

PROOF.

- 1 Suppose not, and let  $\mathbf{x}'$  solve (2.14). Hence,  $\mathbf{p}\mathbf{x}' < \mathbf{p}\mathbf{x}^*$  and  $u(\mathbf{x}') \geq u(\mathbf{x}^*)$ . By local nonsatiation there is a bundle  $\mathbf{x}''$  close enough to  $\mathbf{x}'$  so that  $\mathbf{p}\mathbf{x}'' < \mathbf{p}\mathbf{x}^* = m$  and  $u(\mathbf{x}'') > u(\mathbf{x}^*)$ . But then  $\mathbf{x}^*$  cannot be a solution to (2.13).
- 2 Suppose not, and let  $\mathbf{x}'$  solve (2.13) so that  $u(\mathbf{x}') > u(\mathbf{x}^*)$  and  $\mathbf{p}\mathbf{x}' = \mathbf{p}\mathbf{x}^* = m$ . Since  $\mathbf{p}\mathbf{x}^* > 0$  and utility is continuous, we can find  $0 < t < 1$  such that  $t\mathbf{p}\mathbf{x}' < \mathbf{p}\mathbf{x}^* = m$  and  $u(t\mathbf{x}') > u(\mathbf{x}^*)$ . Hence,  $\mathbf{x}$  cannot solve (2.14).

This proposition leads to four important identities that is summarized in the following proposition.

**Proposition 2.5.7** *Suppose the utility function  $u$  is continuous and locally non-satiated, and suppose that  $m > 0$ . Then we have*

- (1)  $e(\mathbf{p}, v(\mathbf{p}, m)) \equiv m$ . *The minimum expenditure necessary to reach utility  $v(\mathbf{p}, m)$  is  $m$ .*
- (2)  $v(\mathbf{p}, e(\mathbf{p}, u)) \equiv u$ . *The maximum utility from income  $e(\mathbf{p}, u)$  is  $u$ .*
- (3)  $x_i(\mathbf{p}, m) \equiv h_i(\mathbf{p}, v(\mathbf{p}, m))$ . *The Marshallian demand at income  $m$  is the same as the Hicksian demand at utility  $v(\mathbf{p}, m)$ .*

(4)  $h_i(\mathbf{p}, u) \equiv x_i(\mathbf{p}, e(\mathbf{p}, u))$ . The Hicksian demand at utility  $u$  is the same as the Marshallian demand at income  $e(\mathbf{p}, u)$ .

This last identity is perhaps the most important since it ties together the “observable” Marshallian demand function with the “unobservable” Hicksian demand function. Thus, any demanded bundle can be expressed *either* as the solution to the utility maximization problem or the expenditure minimization problem.

A nice application of one of these identities is given in the next proposition:

**Roy’s identity.** *If  $\mathbf{x}(\mathbf{p}, m)$  is the Marshallian demand function, then*

$$x_i(\mathbf{p}, m) = -\frac{\frac{\partial v(\mathbf{p}, m)}{\partial p_i}}{\frac{\partial v(\mathbf{p}, m)}{\partial m}} \text{ for } i = 1, \dots, k$$

*provided that the right-hand side is well defined and that  $p_i > 0$  and  $m > 0$ .*

PROOF. Suppose that  $\mathbf{x}^*$  yields a maximal utility of  $u^*$  at  $(\mathbf{p}^*, m^*)$ . We know from our identities that

$$\mathbf{x}(\mathbf{p}^*, m^*) \equiv \mathbf{h}(\mathbf{p}^*, u^*). \quad (2.15)$$

From another one of the fundamental identities, we also know that

$$u^* \equiv v(\mathbf{p}, e(\mathbf{p}, u^*)).$$

Since this is an identity we can differentiate it with respect to  $p_i$  to get

$$0 = \frac{\partial v(\mathbf{p}^*, m^*)}{\partial p_i} + \frac{\partial v(\mathbf{p}^*, m^*)}{\partial m} \frac{\partial e(\mathbf{p}^*, u^*)}{\partial p_i}.$$

Rearranging, and combining this with identity (2.15), we have

$$x_i(\mathbf{p}^*, m^*) \equiv h_i(\mathbf{p}^*, u^*) \equiv \frac{\partial e(\mathbf{p}^*, u^*)}{\partial p_i} \equiv -\frac{\partial v(\mathbf{p}^*, m^*)/\partial p_i}{\partial v(\mathbf{p}^*, m^*)/\partial m}.$$

Since this identity is satisfied for all  $(\mathbf{p}^*, m^*)$  and since  $\mathbf{x}^* = \mathbf{x}(\mathbf{p}^*, m^*)$ , the result is proved.

**Example 2.5.3 (The General Cobb-Douglas Utility Function)** Consider the indirect Cobb-Douglas utility function:

$$v(p, m) = \prod_{l=1}^L \left( \frac{\alpha_l m}{\alpha p_l} \right)^{\alpha_l}$$

where  $\sum_{l=1}^L \alpha_l$ . Then we have

$$v_{p_l} \equiv \frac{\partial v(p, m)}{\partial p_l} = -\frac{\alpha_l}{p_j} v(p, m)$$

$$v_m \equiv \frac{\partial v(p, m)}{\partial m} = \frac{\alpha}{m} v(p, m)$$

Thus, Roy's identity gives the demand functions as

$$x_l(p, m) = \frac{\alpha_l m}{\alpha p_l}, \quad l = 1, 2, \dots, L.$$

**Example 2.5.4 (The General Leontief Utility Function)** Suppose a preference ordering is represented by the Leontief utility function is given by:

$$u(x) = \min\left\{\frac{x_1}{a_1}, \frac{x_2}{a_2}, \dots, \frac{x_L}{a_L}\right\}$$

which is not itself differentiable, but the indirect utility function:

$$v(p, m) = \frac{m}{ap} \tag{2.16}$$

where  $a = (a_1, a_2, \dots, a_L)$  and  $ap$  is the inner product, which is differentiable. Applying Roy's identity, we have

$$x_l(p, m) = -\frac{v_{p_l}(p, m)}{v_m(p, m)} = \frac{a_l m}{(ap)^2} / \frac{1}{ap} = \frac{a_l m}{ap}$$

Hence, Roy's identity often works well even if the differentiability properties of the statement do not hold.

**Example 2.5.5 (The CES Utility Function)** The CES utility function is given by  $u(x_1, x_2) = (x_1^\rho + x_2^\rho)^{1/\rho}$ . We have derived earlier that the indirect utility function is given by:

$$v(\mathbf{p}, m) = (p_1^r + p_2^r)^{-1/r} m.$$

The demand functions can be found by Roy's law:

$$\begin{aligned} x_l(\mathbf{p}, m) &= \frac{-\partial v(\mathbf{p}, m) / \partial p_l}{\partial v(\mathbf{p}, m) / \partial m} = \frac{\frac{1}{r} (p_1^r + p_2^r)^{-(1+\frac{1}{r})} m r p_l^{r-1}}{(p_1^r + p_2^r)^{-1/r}} \\ &= \frac{p_l^{r-1} m}{(p_1^r + p_2^r)}, \quad l = 1, 2. \end{aligned}$$

## 2.6 Duality Between Direct and Indirect Utility

We have seen how one can recover an indirect utility function from observed demand functions by solving the integrability equations. Here we see how to solve for the direct utility function.

The answer exhibits quite nicely the duality between direct and indirect utility functions. It is most convenient to describe the calculations in terms of the normalized indirect utility function, where we have prices divided by income so that expenditure is identically one. Thus the normalized indirect utility function is given by

$$v(\mathbf{p}) = \max_{\mathbf{x}} u(\mathbf{x})$$

such that  $\mathbf{p}\mathbf{x} = 1$

We then have the following proposition

**Proposition 2.6.1** *Given the indirect utility function  $v(\mathbf{p})$ , the direct utility function can be obtained by solving the following problem:*

$$u(\mathbf{x}) = \min_{\mathbf{p}} v(\mathbf{p})$$

such that  $\mathbf{p}\mathbf{x} = 1$

Proof. Let  $\mathbf{x}$  be the demanded bundle at the prices  $\mathbf{p}$ . Then by definition  $v(\mathbf{p}) = u(\mathbf{x})$ . Let  $\mathbf{p}'$  be any other price vector that satisfies the budget constraint so that  $\mathbf{p}'\mathbf{x} = 1$ . Then since  $\mathbf{x}$  is always a *feasible* choice at the prices  $\mathbf{p}'$ , due to the form of the budget set, the utility-maximizing choice must yield utility at least as great as the utility yielded by  $\mathbf{x}$ ; that is,  $v(\mathbf{p}') \geq u(\mathbf{x}) = v(\mathbf{p})$ . Hence, the minimum of the indirect utility function over all  $\mathbf{p}$ 's that satisfy the budget constraint gives us the utility of  $\mathbf{x}$ .

The argument is depicted in Figure 2.13. Any price vector  $\mathbf{p}$  that satisfies the budget constraint  $\mathbf{p}\mathbf{x} = 1$  must yield a higher utility than  $u(\mathbf{x})$ , which is simply to say that  $u(\mathbf{x})$  solves the minimization problem posed above.

**Example 2.6.1 (Solving for the Direct Utility Function)** Suppose that we have an indirect utility function given by  $v(p_1, p_2) = -a \ln p_1 - b \ln p_2$ . What is its associated direct utility function? We set up the minimization problem:

$$\min_{p_1, p_2} -a \ln p_1 - b \ln p_2$$

such that  $p_1 x_1 + p_2 x_2 = 1$ .

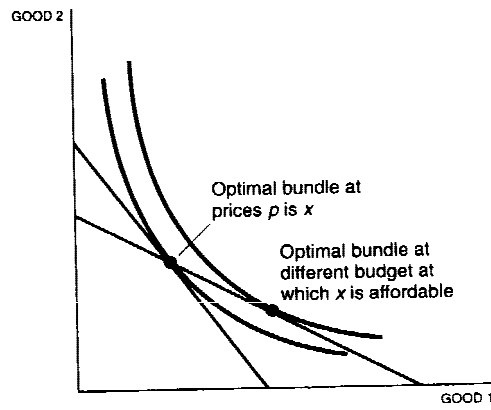


Figure 2.13: Solving for the direct utility function.

The first-order conditions are

$$\begin{aligned} -a/p_1 &= \lambda x_1 \\ -b/p_2 &= \lambda x_2, \end{aligned}$$

or,

$$\begin{aligned} -a &= \lambda p_1 x_1 \\ -b &= \lambda p_2 x_2. \end{aligned}$$

Adding together and using the budget constraint yields

$$\lambda = -a - b.$$

Substitute back into the first-order conditions to find

$$\begin{aligned} p_1 &= \frac{a}{(a+b)x_1} \\ p_2 &= \frac{b}{(a+b)x_2} \end{aligned}$$

These are the choices of  $(p_1, p_2)$  that minimize indirect utility. Now substitute these choices into the indirect utility function:

$$\begin{aligned} u(x_1, x_2) &= -a \ln \frac{a}{(a+b)x_1} - b \ln \frac{b}{(a+b)x_2} \\ &= a \ln x_1 + b \ln x_2 + \text{constant}. \end{aligned}$$

This is the familiar Cobb-Douglas utility function.

The duality between seemingly different ways of representing economic behavior is useful in the study of consumer theory, welfare economics, and many other areas in economics. Many relationships that are difficult to understand when looked at directly become simple, or even trivial, when looked at using the tools of duality.

## 2.7 Properties of Consumer Demand

In this section we will examine the comparative statics of consumer demand behavior: how the consumer's demand changes as prices and income change.

### 2.7.1 Income Changes and Consumption Choice

It is of interest to look at how the consumer's demand changes as we hold prices fixed and allow income to vary; the resulting locus of utility-maximizing bundles is known as the **income expansion path**. From the income expansion path, we can derive a function that relates income to the demand for each commodity (at constant prices). These functions are called **Engel curves**. There are two possibilities: (1) As income increases, the optimal consumption of a good increases. Such a good is called a **normal good**. (2) As income increases, the optimal consumption of a good decreases. Such a good is called **inferior good**.

For the two-good consumer maximization problem, when the income expansion path (and thus each Engel curve) is upper-ward sloping, both goods are normal goods. When the income expansion path could bend backwards, there is one and only one good that is inferior when the utility function is locally non-satiated; increase in income means the consumer actually wants to consume less of the good. (See Figure 2.14)

### 2.7.2 Price Changes and Consumption Choice

We can also hold income fixed and allow prices to vary. If we let  $p_1$  vary and hold  $p_2$  and  $m$  fixed, the locus of tangencies will sweep out a curve known as the **price offer curve**. In the first case in Figure 2.15 we have the ordinary case where a lower price for good 1 leads to greater demand for the good so that the Law of Demand is satisfied; in the

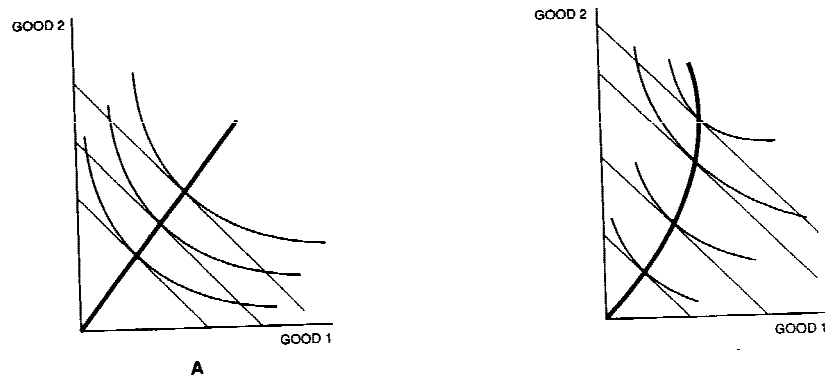


Figure 2.14: Income expansion paths with an interior good.

second case we have a situation where a decrease in the price of good 1 brings about a *decreased* demand for good 1. Such a good is called a **Giffen good**.

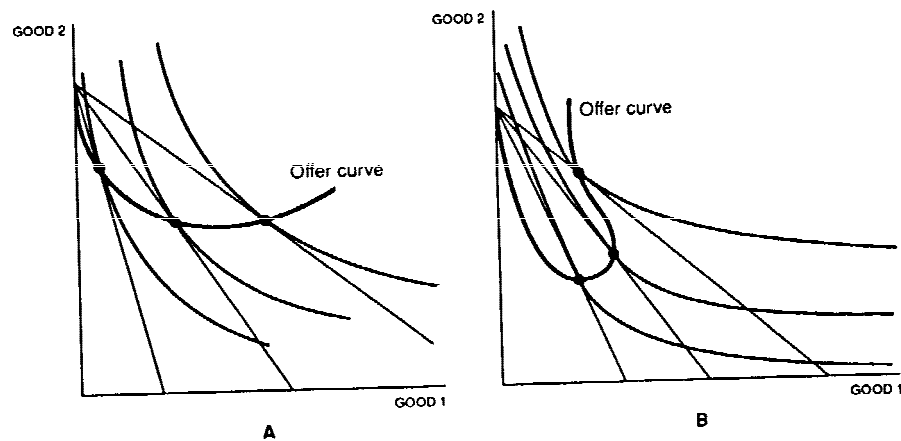


Figure 2.15: Offer curves. In panel A the demand for good 1 increases as the price decreases so it is an ordinary good. In panel B the demand for good 1 decreases as its price decreases, so it is a Giffen good.

### 2.7.3 Income-Substitution Effect: The Slutsky Equation

In the above we see that a fall in the price of a good may have two sorts of effects: **substitution effect**—one commodity will become less expensive than another, and **income effect** — total “purchasing power” increases. A fundamental result of the theory of the consumer, the Slutsky equation, relates these two effects.

Even though the compensated demand function is not *directly* observable, we shall see that its derivative can be easily calculated from observable things, namely, the derivative of the Marshallian demand with respect to price and income. This relationship is known as the **Slutsky equation**.

**Slutsky equation.**

$$\frac{\partial x_j(\mathbf{p}, m)}{\partial p_i} = \frac{\partial h_j(\mathbf{p}, v(\mathbf{p}, m))}{\partial p_i} - \frac{\partial x_j(\mathbf{p}, m)}{\partial m} x_i(\mathbf{p}, m)$$

PROOF. Let  $\mathbf{x}^*$  maximize utility at  $(\mathbf{p}^*, m)$  and let  $u^* = u(\mathbf{x}^*)$ . It is identically true that

$$h_j(\mathbf{p}^*, u^*) \equiv x_j(\mathbf{p}, e(\mathbf{p}, u^*)).$$

We can differentiate this with respect to  $p_i$  and evaluate the derivative at  $\mathbf{p}^*$  to get

$$\frac{\partial h_j(\mathbf{p}^*, u^*)}{\partial p_i} = \frac{\partial x_j(\mathbf{p}^*, m^*)}{\partial p_i} + \frac{\partial x_j(\mathbf{p}^*, m^*)}{\partial m} \frac{\partial e(\mathbf{p}^*, u^*)}{\partial p_i}.$$

Note carefully the meaning of this expression. The left-hand side is how the compensated demand changes when  $p_i$  changes. The right-hand side says that this change is equal to the change in demand holding expenditure fixed at  $m^*$  *plus* the change in demand when income changes *times* how much income has to change to keep utility constant. But this last term,  $\partial e(\mathbf{p}^*, u^*)/\partial p_i$ , is just  $x_i^*$ ; rearranging gives us

$$\frac{\partial x_j(\mathbf{p}^*, m^*)}{\partial p_i} = \frac{\partial h_j(\mathbf{p}^*, u^*)}{\partial p_i} - \frac{\partial x_j(\mathbf{p}^*, m^*)}{\partial m} x_i^*$$

which is the Slutsky equation.

There are other ways to derive Slutsky's equations that can be found in Varian (1992).

The Slutsky equation decomposes the demand change induced by a price change  $\Delta p_i$  into two separate effects: the **substitution effect** and the **income effect**:

$$\Delta x_j \approx \frac{\partial x_j(\mathbf{p}, m)}{\partial p_i} \Delta p_i = \frac{\partial h_j(\mathbf{p}, u)}{\partial p_i} \Delta p_i - \frac{\partial x_j(\mathbf{p}, m)}{\partial m} x_i^* \Delta p_i$$

As we mentioned previously, the restrictions all about the Hicksian demand functions are not directly observable. However, as indicated by the Slutsky equation, we can express the derivatives of  $\mathbf{h}$  with respect to  $\mathbf{p}$  as derivatives of  $\mathbf{x}$  with respect to  $\mathbf{p}$  and  $m$ , and these are observable.

Also, Slutsky's equation and the negative sime-definite matrix on Hicksian demand functions given in Proposition 2.5.3 give us the following result on the Marshallian demand functions:



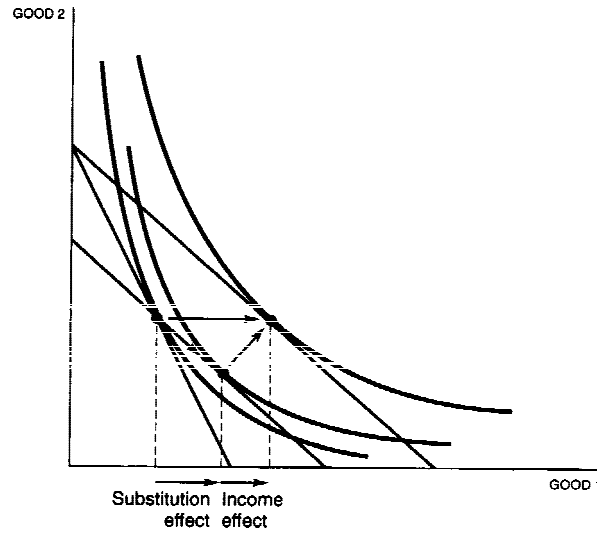


Figure 2.16: The Hicks decomposition of a demand change into two effects: the substitution effect and the income effect.

**Proposition 2.7.1** *The substitution matrix  $\left( \frac{\partial x_j(\mathbf{p}, m)}{\partial p_i} + \frac{\partial x_j(\mathbf{p}, u)}{\partial m} x_i \right)$  is a symmetric, negative semi-definite matrix.*

This is a rather nonintuitive result: a particular combination of price and income derivatives has to result in a negative semidefinite matrix.

**Example 2.7.1 (The Cobb-Douglas Slutsky equation)** Let us check the Slutsky equation in the Cobb-Douglas case. As we've seen, in this case we have

$$\begin{aligned} v(p_1, p_2, m) &= mp_1^{-\alpha} p_2^{\alpha-1} \\ e(p_1, p_2, u) &= up_1^\alpha p_2^{1-\alpha} \\ x_1(p_1, p_2, m) &= \frac{\alpha m}{p_1} \\ h_1(p_1, p_2, u) &= \alpha p_1^{\alpha-1} p_2^{1-\alpha} u. \end{aligned}$$

Thus

$$\begin{aligned}
\frac{\partial x_1(\mathbf{p}, m)}{\partial p_1} &= -\frac{\alpha m}{p_1^2} \\
\frac{\partial x_1(\mathbf{p}, m)}{\partial m} &= -\frac{\alpha}{p_1} \\
\frac{\partial h_1(\mathbf{p}, u)}{\partial p_1} &= \alpha(\alpha - 1)p_1^{\alpha-2}p_2^{1-\alpha}u \\
\frac{\partial h_1(\mathbf{p}, v(\mathbf{p}, m))}{\partial p_1} &= \alpha(\alpha - 1)p_1^{\alpha-2}p_2^{1-\alpha}mp_1^{-\alpha}p_2^{\alpha-1} \\
&= \alpha(\alpha - 1)p_1^{-2}m.
\end{aligned}$$

Now plug into the Slutsky equation to find

$$\begin{aligned}
\frac{\partial h_1}{\partial p_1} - \frac{\partial x_1}{\partial m}x_1 &= \frac{\alpha(\alpha - 1)m}{p_1^2} - \frac{\alpha}{p_1} \frac{\alpha m}{p_1} \\
&= \frac{[\alpha(\alpha - 1) - \alpha^2]m}{p_1^2} \\
&= \frac{-\alpha m}{p_1^2} = \frac{\partial x_1}{\partial p_1}.
\end{aligned}$$

#### 2.7.4 Continuity and Differentiability of Demand Functions

Up until now we have assumed that the demand functions are nicely behaved; that is, that they are continuous and even differentiable functions. Are these assumptions justifiable?

**Proposition 2.7.2 (Continuity of Demand Function)** *Suppose  $\succeq$  is continuous and weakly convex, and  $(p, m) > 0$ . Then,  $x(p, m)$  is a upper hemi-continuous convex-valued correspondence. Furthermore, if the weak convexity is replaced by the strict convexity,  $x(p, m)$  is a continuous single-valued function.*

*Proof.* First note that, since  $(p, m) > 0$ , one can show that the budget constrained set  $B(p, m)$  is a continuous correspondence with non-empty and compact values and  $\succeq_i$  is continuous. Then, by the Maximum Theorem, we know the demand correspondence  $x(p, m)$  is upper hemi-continuous. We now show  $x(p, m)$  is convex. Suppose  $x$  and  $x'$  are two optimal consumption bundles. Let  $x_t = tx + (1 - t)x'$  for  $t \in [0, 1]$ . Then,  $x_t$  also satisfied the budget constraint, and by weak convexity of  $\succeq$ , we have  $x_t = tx + (1 - t)x' \succeq x$ . Because  $x$  is an optimal consumption bundle, we must have  $x_t \sim x$  and thus  $x_t$  is also an optimal consumption bundle.

Now, when  $\succ$  is strictly convex,  $x(p, m)$  is then single-valued by Proposition 2.4.2, and thus it is a continuous function since a upper hemi-continuous correspondence is a continuous function when it is single-valued.

A demand correspondence may not be continuous for non-convex preference ordering, as illustrated in Figure 2.17. Note that, in the case depicted in Figure 2.17, a small change in the price brings about a large change in the demanded bundles: the demand correspondence is discontinuous.

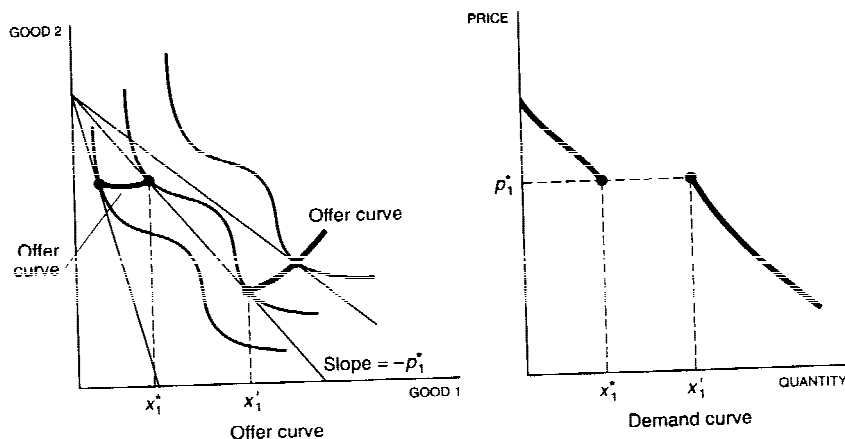


Figure 2.17: Discontinuous demand. Demand is discontinuous due to non-convex preferences

Sometimes, we need to consider the slopes of demand curves and hence we would like a demand function is differentiable. What conditions can guarantee the differentiability? We give the following proposition without proof.

**Proposition 2.7.3** *Suppose  $x > 0$  solves the consumer's utility maximization problem at  $(p, m) > 0$ . If*

- (1)  *$u$  is twice continuously differentiable on  $\mathbb{R}_{++}^L$ ,*
- (2)  *$\frac{\partial u(x)}{\partial x_l} > 0$  for some  $l = 1, \dots, L$ ,*
- (3) *the bordered Hessian of  $u$  has nonzero determinant at  $x$ ,*

*then  $x(p, m)$  is differentiable at  $(p, m)$ .*

### 2.7.5 Inverse Demand Functions

In many applications it is of interest to express demand behavior by describing prices as a function of quantities. That is, given some vector of goods  $\mathbf{x}$ , we would like to find a vector of prices  $\mathbf{p}$  and an income  $m$  at which  $\mathbf{x}$  would be the demanded bundle.

Since demand functions are homogeneous of degree zero, we can fix income at some given level, and simply determine prices relative to this income level. The most convenient choice is to fix  $m = 1$ . In this case the first-order conditions for the utility maximization problem are simply

$$\begin{aligned}\frac{\partial u(\mathbf{x})}{\partial x_i} - \lambda p_i &= 0 \quad \text{for } i, \dots, k \\ \sum_{i=1}^k p_i x_i &= 1.\end{aligned}$$

We want to eliminate  $\lambda$  from this set of equations.

To do so, multiply each of the first set of equalities by  $x_i$  and sum them over the number of goods to get

$$\sum_{i=1}^k \frac{\partial u(\mathbf{x})}{\partial x_i} x_i = \lambda \sum_{i=1}^k p_i x_i = \lambda.$$

Substitute the value of  $\lambda$  back into the first expression to find  $\mathbf{p}$  as function of  $\mathbf{x}$ :

$$p_i(\mathbf{x}) = \frac{\frac{\partial u(\mathbf{x})}{\partial x_i}}{\sum_{j=1}^k \frac{\partial u(\mathbf{x})}{\partial x_j} x_j}. \quad (2.17)$$

Given any vector of demands  $\mathbf{x}$ , we can use this expression to find the price vector  $\mathbf{p}(\mathbf{x})$  which will satisfy the necessary conditions for maximization. If the utility function is quasi-concave so that these necessary conditions are indeed sufficient for maximization, then this will give us the inverse demand relationship.

What happens if the utility function is not everywhere quasi-concave? Then there may be some bundles of goods that will not be demanded at any price; any bundle on a non-convex part of an indifference curve will be such a bundle.

There is a dual version of the above formula for inverse demands that can be obtained from the duality between direct utility function and indirect utility function we discussed earlier. The argument given there shows that the demanded bundle  $\mathbf{x}$  must minimize

indirect utility over all prices that satisfy the budget constraint. Thus  $\mathbf{x}$  must satisfy the first-order conditions

$$\begin{aligned}\frac{\partial v(\mathbf{p})}{\partial p_l} - \mu x_l &= 0 \quad \text{for } l = 1, \dots, L \\ \sum_{i=1}^L p_i x_i &= 1.\end{aligned}$$

Now multiply each of the first equations by  $p_l$  and sum them to find that  $\mu = \sum_{l=1}^L \frac{\partial v(\mathbf{p})}{\partial p_l} p_l$ . Substituting this back into the first-order conditions, we have an expression for the demanded bundle as a function of the normalized indirect utility function:

$$x_i(\mathbf{p}) = \frac{\frac{\partial v(\mathbf{p})}{\partial p_i}}{\sum_{j=1}^L \frac{\partial v(\mathbf{p})}{\partial p_j} p_j}. \quad (2.18)$$

Note the nice duality: the expression for the direct demand function, (2.18), and the expression for the inverse demand function (2.17) have the same form. This expression can also be derived from the definition of the normalized indirect utility function and Roy's identity.

## 2.8 The Integrability Problem

Given a system of demand functions  $x(p, m)$ . Is there necessarily a utility function from which these demand functions can be derived? This question is known as the **integrability problem**. We will show how to solve this problem by solving a differential equation and integrating back, ultimately to the utility function. The Slutsky matrix plays a key role in this process.

We have seen that the utility maximization hypothesis imposes certain observable restrictions on consumer behavior. If a demand function  $x(p, m)$  is well-behaved, our previous analysis has shown that  $x(p, m)$  satisfies the following five conditions:

1. Nonnegativity:  $x(p, m) \geq 0$ .
2. Homogeneity:  $x(tp, tm) = x(p, m)$ .
3. Budget Balancedness:  $px(p, m) = m$ .
4. Symmetry: The Slutsky matrix  $S \equiv \left( \frac{\partial x_i(\mathbf{p}, m)}{\partial p_j} + \frac{\partial x_i(\mathbf{p}, m)}{\partial m} x_j(\mathbf{p}, m) \right)$  is symmetric.

5. Negative Semi-definite: The matrix  $S$  is negative semi-definite.

The main result of the integrability problem is that these conditions, together with some technical assumptions, are in fact sufficient as well as necessary for the integrability process as shown by Hurwicz and Uzawa (1971). This result is very important from the point view of political economy. The utility maximization approach to the study of consumer behavior sometimes are criticized because they think the notion of utility is a psychological measurement and cannot be observed, and thus, they think the demand function from utility maximization is meaningless. The integrability result, however, tells us that a utility function can be derived from observable data on demand although the utility function is not directly observable. This impress result warrants of a formal statement.

**Theorem 2.8.1** *A continuous differentiable function  $x : \mathbb{R}_{++}^{L+1} \rightarrow \mathbb{R}_+^L$  is the demand function generalized by some increasing, quasi-concave utility function  $u$  if (and only if, when  $u$  is continuous, strictly increasing and strictly quasi-concave) it satisfy homogeneity, budget balancedness, symmetry, and negative semi-definiteness.*

The proof of the theorem is somehow complicated and can be found in Hurwicz and Uzawa (1971). So it is omitted here.

To actually find a utility function from a give system of demand functions, we must find an equation to integrate. As it turns out, it is somewhat easier to deal with the integrability problem in terms of the expenditure function rather than the indirect utility function.

Recall that from Shephard's lemma given in Proposition 2.5.2,

$$\frac{\partial e(\mathbf{p}, u)}{\partial p_i} = x_i(p, m) = x_i(\mathbf{p}, e(p, u)) \quad i = 1, \dots, L. \quad (2.19)$$

We also specify a boundary condition of the form  $e(p^*, u) = c$  where  $p^*$  and  $c$  are given. The system of equations given in (2.19) is a system of **partial differential** equations. It is well-known that a system of partial differential equations of the form

$$\frac{\partial f(\mathbf{p})}{\partial p_i} = g_i(\mathbf{p}) \quad i = 1, \dots, k$$

has a (local) solution if and only if

$$\frac{\partial g_i(\mathbf{p})}{\partial p_j} = \frac{\partial g_j(\mathbf{p})}{\partial p_i} \quad \text{all } i \text{ and } j.$$

Applying this condition to the above problem, we see that it reduces to requiring that the matrix

$$\left( \frac{\partial x_i(\mathbf{p}, m)}{\partial p_j} + \frac{\partial x_i(\mathbf{p}, m)}{\partial m} \frac{\partial e(\mathbf{p}, u)}{\partial p_j} \right)$$

is symmetric. But this is just the Slutsky restriction! Thus the Slutsky restrictions imply that the demand functions can be “integrated” to find an expenditure function consistent with the observed choice behavior.

Under the assumption that all five of the properties listed at the beginning of this section hold, the solution function  $e$  will be an expenditure function. Inverting the found expenditure function, we can find the indirect utility function. Then using the duality between the direct utility function and indirect utility function we will study in the next section, we can determine the direct utility function.

**Example 2.8.1 (The General Cobb-Douglas Utility Function)** Consider the demand functions

$$x_i(p, m) = \frac{a_i m}{p_i} = \frac{\alpha_i m}{\alpha p_i}$$

where  $\alpha = \sum_{l=1}^L \alpha_l$ . The system (2.19) becomes

$$\frac{\partial e(\mathbf{p}, u)}{\partial p_i} = \frac{\alpha_i m}{\alpha p_i} \quad i = 1, \dots, L. \quad (2.20)$$

The  $i$ -equation can be integrated with respect to  $p_i$  to obtain.

$$\ln e(p, u) = \frac{\alpha_i}{\alpha} \ln p_i + c_i$$

where  $c_i$  does not depend on  $p_i$ , but it may depend on  $p_j$  for  $j \neq i$ . Thus, combining these equations we find

$$\ln e(p, u) = \sum_{i=1}^L \frac{\alpha_i}{\alpha} \ln p_i + c$$

where  $c$  is independent of all  $p_i$ 's. The constant  $c$  represents the freedom that we have in setting the boundary condition. For each  $u$ , let us take  $p^* = (1, \dots, 1)$  and use the boundary condition  $e(p^*, u) = u$ . Then it follows that

$$\ln e(p, u) = \sum_{i=1}^L \frac{\alpha_i}{\alpha} \ln p_i + \ln u.$$

Inverting the above equation, we have

$$\ln v(p, m) = - \sum_{i=1}^L \frac{\alpha_i}{\alpha} \ln p_i + \ln m$$

which is a monotonic transformation of the indirect utility function for a Cobb-Douglas we found previously.

## 2.9 Revealed Preference

### 2.9.1 Axioms of Revealed Preferences

The basic preference axioms sometimes are criticized as being too strong on the grounds that individuals are unlikely to make choices through conscious use of a preference relation. One response to this criticism is to develop an alternative theory on the basis of a weaker set of hypotheses. One of the most interesting alternative theories is that of revealed preference, which is discussed in this section.

The basic principle of revealed preference theory is that preference statements should be constructed only from observable decisions, that is, from actual choice made by a consumer. An individual preference relation, even if it exists, can never be directly observed in the market. The best that we may hope for in practice is a list of the choices made under different circumstances. For example, we may have some observations on consumer behavior that take the form of a list of prices,  $\mathbf{p}^t$ , and the associated chosen consumption bundles,  $\mathbf{x}^t$  for  $t = 1, \dots, T$ . How can we tell whether these data could have been generated by a utility-maximizing consumer? Revealed preference theory focuses on the choices made by a consumer, not on a hidden preference relation.

We will say that a utility function rationalizes the observed behavior  $(\mathbf{p}^t \mathbf{x}^t)$  for  $t = 1, \dots, T$  if  $u(\mathbf{x}^t) \geq u(\mathbf{x})$  for all  $\mathbf{x}$  such that  $\mathbf{p}^t \mathbf{x}^t \geq \mathbf{p}^t \mathbf{x}$ . That is,  $u(\mathbf{x})$  rationalizes the observed behavior if it achieves its maximum value on the budget set at the chosen bundles. Suppose that the data were generated by such a maximization process. What observable restrictions must the observed choices satisfy?

Without any assumptions about  $u(\mathbf{x})$  there is a trivial answer to this question, namely, no restrictions. For suppose that  $u(\mathbf{x})$  were a constant function, so that the consumer was indifferent to all observed consumption bundles. Then there would be no restrictions imposed on the patterns of observed choices: anything is possible.

To make the problem interesting, we have to rule out this trivial case. The easiest way to do this is to require the underlying utility function to be locally non-satiated. Our



question now becomes: what are the observable restrictions imposed by the maximization of a locally non-satiated utility function?

**Direct Revealed Preference:** If  $\mathbf{p}^t \mathbf{x}^t \geq \mathbf{p}^t \mathbf{x}$ , then  $u(\mathbf{x}^t) \geq u(\mathbf{x})$ . We will say that  $\mathbf{x}^t$  is directly revealed preferred to  $\mathbf{x}$ , and write  $\mathbf{x}^t R^D \mathbf{x}$ .

This condition means that if  $\mathbf{x}^t$  was chosen when  $\mathbf{x}$  could have been chosen, the utility of  $\mathbf{x}^t$  must be at least as large as the utility of  $\mathbf{x}$ . As a consequence of this definition and the assumption that the data were generated by utility maximization, we can conclude that “ $\mathbf{x}^t R^D \mathbf{x}$  implies  $u(\mathbf{x}^t) \geq u(\mathbf{x})$ .”

**Strictly Direct Revealed Preference:** If  $\mathbf{p}^t \mathbf{x}^t > \mathbf{p}^t \mathbf{x}$ , then  $u(\mathbf{x}^t) > u(\mathbf{x})$ . We will say that  $\mathbf{x}^t$  is strictly directly revealed preferred to  $\mathbf{x}$  and write  $\mathbf{x}^t P^D \mathbf{x}$ .

It is not hard to show that local non-satiation implies this conclusion. For we know from the previous paragraph that  $u(\mathbf{x}^t) \geq u(\mathbf{x})$ ; if  $u(\mathbf{x}^t) = u(\mathbf{x})$ , then by local non-satiation there would exist some other  $\mathbf{x}'$  close enough to  $\mathbf{x}$  so that  $\mathbf{p}^t \mathbf{x}^t > \mathbf{p}^t \mathbf{x}'$  and  $u(\mathbf{x}') > u(\mathbf{x}) = u(\mathbf{x}^t)$ . This contradicts the hypothesis of utility maximization.

**Revealed Preference:**  $\mathbf{x}^t$  is said to be revealed preferred to  $\mathbf{x}$  if there exists a finite number of bundles  $x_1, x_2, \dots, x_n$  such that  $\mathbf{x}^t R^D \mathbf{x}^1, \mathbf{x}^1 R^D \mathbf{x}^2, \dots, \mathbf{x}^n R^D \mathbf{x}$ . In this case, we write  $\mathbf{x}^t R \mathbf{x}$ .

The relation  $R$  constructed above by considering chains of  $R^D$  is sometimes called the **transitive closure** of the relation  $R^D$ . If we assume that the data were generated by utility maximization, it follows that “ $\mathbf{x}^t R \mathbf{x}$  implies  $u(\mathbf{x}^t) \geq u(\mathbf{x})$ .”

Consider two observations  $\mathbf{x}^t$  and  $\mathbf{x}^s$ . We now have a way to determine whether  $u(\mathbf{x}^t) \geq u(\mathbf{x}^s)$  and an observable condition to determine whether  $u(\mathbf{x}^s) > u(\mathbf{x}^t)$ . Obviously, these two conditions should not both be satisfied. This condition can be stated as the

**GENERALIZED AXIOM OF REVEALED PREFERENCE (GARP):** *If  $\mathbf{x}^t$  is revealed preferred to  $\mathbf{x}^s$ , then  $\mathbf{x}^s$  cannot be strictly directly revealed preferred to  $\mathbf{x}^t$ .*

Using the symbols defined above, we can also write this axiom as

**GARP:**  $\mathbf{x}^t R \mathbf{x}^s$  implies not  $\mathbf{x}^s P^D \mathbf{x}^t$ . In other words,  $\mathbf{x}^t R \mathbf{x}^s$  implies  $p^s \mathbf{x}^s \leq p^s \mathbf{x}^t$ .

As the name implies, GARP is a generalization of various other revealed preference tests. Here are two standard conditions.

**WEAK AXIOM OF REVEALED PREFERENCE (WARP):** *If  $\mathbf{x}^t R^D \mathbf{x}^s$  and  $\mathbf{x}^t$  is not equal to  $\mathbf{x}^s$ , then it is not the case that  $\mathbf{x}^s R^D \mathbf{x}^t$ , i.e.,  $p_t x_t \geq p_t x_s$  implies  $p_s x_t > p_s x_s$ .*

**STRONG AXIOM OF REVEALED PREFERENCE (SARP):** If  $\mathbf{x}^t R \mathbf{x}^s$  and  $\mathbf{x}^t$  is not equal to  $\mathbf{x}^s$ , then it is not the case that  $\mathbf{x}^s R \mathbf{x}^t$ .

Each of these axioms requires that there be a *unique* demand bundle at each budget, while GARP allows for multiple demanded bundles. Thus, GARP allows for flat spots in the indifference curves that generated the observed choices.

## 2.9.2 Characterization of Revealed Preference Maximization

If the data  $(\mathbf{p}^t, \mathbf{x}^t)$  were generated by a utility-maximizing consumer with nonsatiated preferences, the data must satisfy GARP. Hence, GARP is an observable consequence of utility maximization. But does it express all the implications of that model? If some data satisfy this axiom, is it necessarily true that it must come from utility maximization, or at least be thought of in that way? Is GARP a *sufficient* condition for utility maximization?

It turns out that it is. If a finite set of data is consistent with GARP, then there exists a utility function that rationalizes the observed behavior — i.e., there exists a utility function that could have generated that behavior. Hence, GARP exhausts the list of restrictions imposed by the maximization model.

We state the following theorem without proof.

**Afriat's theorem.** Let  $(\mathbf{p}^t, \mathbf{x}^t)$  for  $t = 1, \dots, T$  be a finite number of observations of price vectors and consumption bundles. Then the following conditions are equivalent.

- (1) There exists a locally nonsatiated utility function that rationalizes the data;
- (2) The data satisfy GARP;
- (3) There exist positive numbers  $(u^t, \lambda^t)$  for  $t = 1, \dots, T$  that satisfy the Afriat inequalities:

$$u^s \leq u^t + \lambda^t \mathbf{p}^t(\mathbf{x}^s - \mathbf{x}^t) \quad \text{for all } t, s;$$

- (4) There exists a locally nonsatiated, continuous, concave, monotonic utility function that rationalizes the data.

Thus, Afriat's theorem states that a finite set of observed price and quantity data satisfy GARP if and only if there exists a locally non-satiated, continuous, increasing, and concave utility function that rationalizes the data.

Condition (3) in Afriat's theorem has a natural interpretation. Suppose that  $u(\mathbf{x})$  is a concave, differentiable utility function that rationalizes the observed choices. The fact that  $u(\mathbf{x})$  is differentiable implies it must satisfy the  $T$  first-order conditions

$$\mathbf{D}u(\mathbf{x}^t) = \lambda^t \mathbf{p}^t \quad (2.21)$$

The fact that  $u(\mathbf{x})$  is concave implies that it must satisfy the concavity conditions

$$u(\mathbf{x}^t) \leq u(\mathbf{x}^s) + \mathbf{D}u(\mathbf{x}^s)(\mathbf{x}^t - \mathbf{x}^s). \quad (2.22)$$

Substituting from (2.21) into (2.22), we have

$$u(\mathbf{x}^t) \leq u(\mathbf{x}^s) + \lambda^s \mathbf{p}^s(\mathbf{x}^t - \mathbf{x}^s).$$

Hence, the Afriat numbers  $u^t$  and  $\lambda^t$  can be interpreted as utility levels and marginal utilities that are consistent with the observed choices.

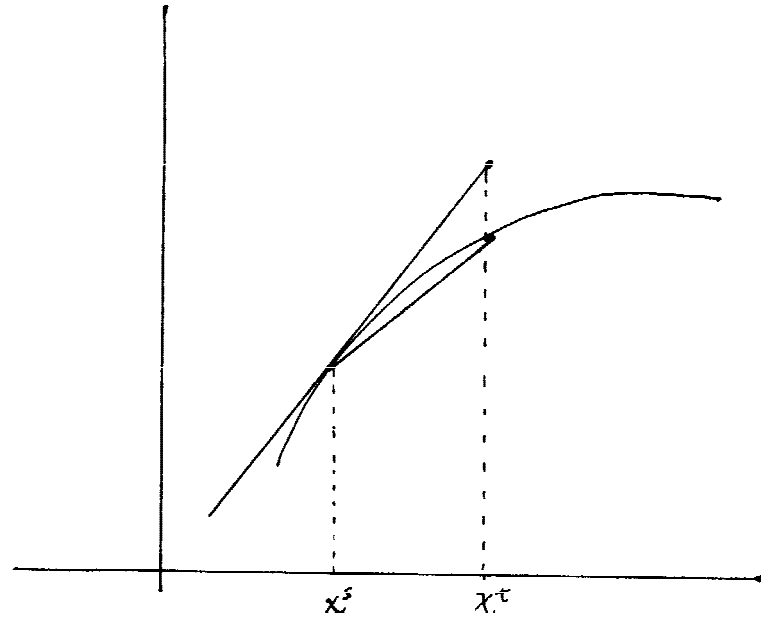


Figure 2.18: Concave function.

The reason the inequality holds for a concave function is because that, from Figure 2.18, we have

$$\frac{u(x^t) - u(x^s)}{x^t - x^s} \leq u'(x^s). \quad (2.23)$$

Thus, we have  $u(x^t) \leq u(x^s) + u'(x^s)(x^t - x^s)$ .

The most remarkable implication of Afriat's theorem is that (1) implies (4): if there is any locally nonsatiated utility function at all that rationalizes the data, there must exist a continuous, monotonic, and concave utility function that rationalizes the data. If the underlying utility function had the “wrong” curvature at some points, we would never observe choices being made at such points because they wouldn't satisfy the right second-order conditions. Hence market data do not allow us to reject the hypotheses of convexity and monotonicity of preferences.

## 2.10 Recoverability

Since the revealed preference conditions are a complete set of the restrictions imposed by utility-maximizing behavior, they must contain all of the information available about the underlying preferences. It is more-or-less obvious now to use the revealed preference relations to determine the preferences among the *observed* choices,  $\mathbf{x}^t$ , for  $t = 1, \dots, T$ . However, it is less obvious to use the revealed preference relations to tell you about preference relations between choices that have never been observed.

This is easiest to see using an example, Figure 2.19 depicts a single observation of choice behavior,  $(\mathbf{p}^1, \mathbf{x}^1)$ . What does this choice imply about the indifference curve through a bundle  $\mathbf{x}^0$ ? Note that  $\mathbf{x}^0$  has not been previously observed; in particular, we have no data about the prices at which  $\mathbf{x}^0$  would be an optimal choice.

Let's try to use revealed preference to “bound” the indifference curve through  $\mathbf{x}^0$ . First, we observe that  $\mathbf{x}^1$  is revealed preferred to  $\mathbf{x}^0$ . Assume that preferences are convex and monotonic. Then all the bundles on the line segment connecting  $\mathbf{x}^0$  and  $\mathbf{x}^1$  must be at least as good as  $\mathbf{x}^0$ , and all the bundles that lie to the northeast of this bundle are at least as good as  $\mathbf{x}^0$ . Call this set of bundles  $RP(\mathbf{x}^0)$ , for “revealed preferred” to  $\mathbf{x}^0$ . It is not difficult to show that this is the best “inner bound” to the upper contour set through the point  $\mathbf{x}^0$ .

To derive the best outer bound, we must consider all possible budget lines passing through  $\mathbf{x}^0$ . Let  $RW$  be the set of all bundles that are revealed worse than  $\mathbf{x}^0$  for all these budget lines. The bundles in  $RW$  are certain to be worse than  $\mathbf{x}^0$  no matter what budget line is used.

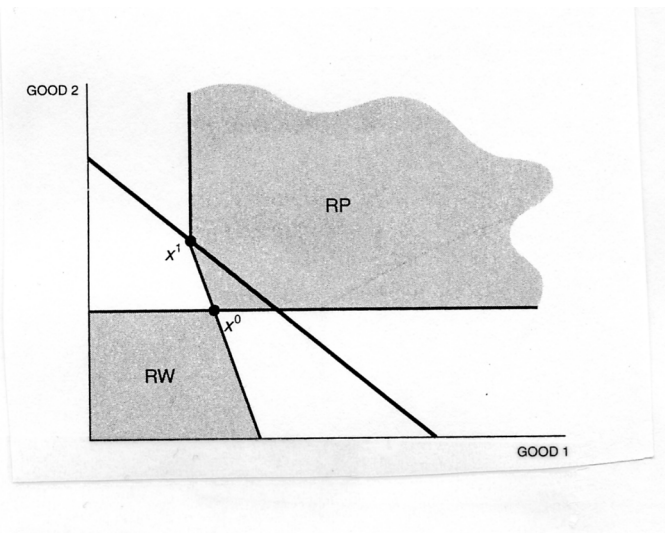


Figure 2.19: Inner and outer bounds.  $RP$  is the inner bound to the indifference curve through  $x^0$ ; the consumption of  $RW$  is the outer bound.

The outer bound to the upper contour set at  $\mathbf{x}^0$  is then defined to be the complement of this set:  $NRW = \text{all bundles not in } RW$ . This is the best outer bound in the sense that any bundle not in this set cannot ever be revealed preferred to  $\mathbf{x}^0$  by a consistent utility-maximizing consumer. Why? Because by construction, a bundle that is not in  $NRW(\mathbf{x}^0)$  must be in  $RW(\mathbf{x}^0)$  in which case it would be revealed *worse* than  $\mathbf{x}^0$ .

In the case of a single observed choice, the bounds are not very tight. But with many choices, the bounds can become quite close together, effectively trapping the true indifference curve between them. See Figure 2.20 for an illustrative example. It is worth tracing through the construction of these bounds to make sure that you understand where they come from. Once we have constructed the inner and outer bounds for the upper contour sets, we have recovered essentially all the information about preferences that is not aimed in the observed demand behavior. Hence, the construction of  $RP$  and  $RW$  is analogous to solving the integrability equations.

Our construction of  $RP$  and  $RW$  up until this point has been graphical. However, it is possible to generalize this analysis to multiple goods. It turns out that determining whether one bundle is revealed preferred or revealed worse than another involves checking to see whether a solution exists to a particular set of linear inequalities.

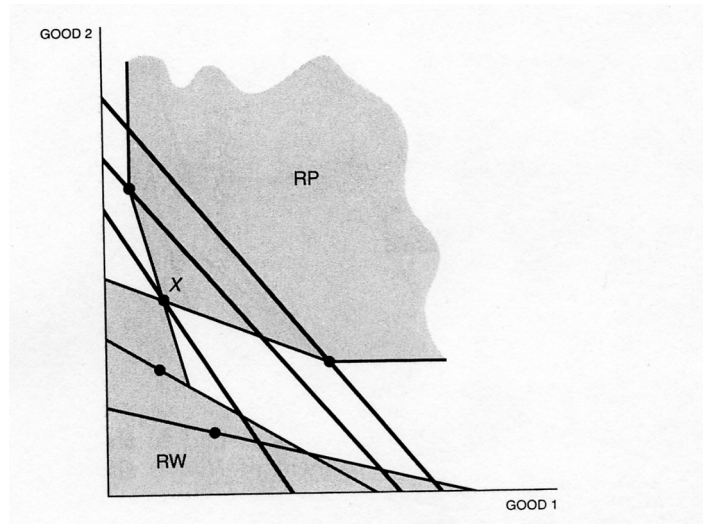


Figure 2.20: Inner and outer bounds. When there are several observations, the inner bound and outer bound can be quite tight.

## 2.11 Topics in Demand Behavior

In this section we investigate several topics in demand behavior. Most of these have to do with special forms of the budget constraint or preferences that lead to special forms of demand behavior. There are many circumstances where such special cases are very convenient for analysis, and it is useful to understand how they work.

### 2.11.1 Endowments in the Budget Constraint

In our study of consumer behavior we have taken income to be exogenous. But in more elaborate models of consumer behavior it is necessary to consider how income is generated. The standard way to do this is to think of the consumer as having some **endowment**  $\omega = (\omega_1, \dots, \omega_L)$  of various goods which can be sold at the current market prices  $\mathbf{p}$ . This gives the consumer income  $m = \mathbf{p}\omega$  which can be used to purchase other goods.

The utility maximization problem becomes

$$\begin{aligned} \max_{\mathbf{x}} u(\mathbf{x}) \\ \text{such that } \mathbf{p}\mathbf{x} &= \mathbf{p}\omega. \end{aligned}$$

This can be solved by the standard techniques to find a demand function  $\mathbf{x}(\mathbf{p}, \mathbf{p}\omega)$ . The **net demand** for good  $i$  is  $x_i - \omega_i$ . The consumer may have positive or negative net

demands depending on whether he wants more or less of something than is available in his endowment.

In this model prices influence the value of what the consumer has to sell as well as the value of what the consumer wants to sell. This shows up most clearly in Slutsky's equation, which we now derive. First, differentiate demand with respect to price:

$$\frac{dx_i(\mathbf{p}, \mathbf{p}\omega)}{dp_j} = \left. \frac{\partial x_i(\mathbf{p}, \mathbf{p}\omega)}{\partial p_j} \right|_{\mathbf{p}\omega=\text{constant}} + \frac{\partial x_i(\mathbf{p}, \mathbf{p}\omega)}{\partial m} \omega_j.$$

The first term in the right-hand side of this expression is the derivative of demand with respect to price, holding income fixed. The second term is the derivative of demand with respect to income, times the change in income. The first term can be expanded using Slutsky's equation. Collecting terms we have

$$\frac{dx_i(\mathbf{p}, \mathbf{p}\omega)}{dp_j} = \frac{\partial h_i(\mathbf{p}, u)}{\partial p_j} + \frac{\partial x_i(\mathbf{p}, \mathbf{p}\omega)}{\partial m} (\omega_j - x_j).$$

Now the income effect depends on the net demand for good  $j$  rather than the gross demand.

## 2.11.2 Income-Leisure Choice Model

Suppose that a consumer chooses two goods, consumption and "leisure". Let  $\ell$  be the number of hours and  $\bar{L}$  be the maximum number of hours that the consumer can work. We then have  $L = \bar{L} - \ell$ . She also has some nonlabor income  $m$ . Let  $u(c, L)$  be the utility of consumption and leisure and write the utility maximization problem as

$$\begin{aligned} \max_{c, L} u(c, L) \\ \text{such that } pc + wL = w\bar{L} + m. \end{aligned}$$

This is essentially the same form that we have seen before. Here the consumer "sells" her endowment of labor at the price  $w$  and then buys some back as leisure.

Slutsky's equation allows us to calculate how the demand for leisure changes as the wage rate changes. We have

$$\frac{dL(p, w, m)}{dw} = \frac{\partial L(p, w, u)}{\partial w} + \frac{\partial L(p, w, m)}{\partial m} [\bar{L} - L].$$

Note that the term in brackets is nonnegative by definition, and almost surely positive in practice. This means that the derivative of leisure demand is the sum of a negative

number and a positive number and is inherently ambiguous in sign. In other words, an increase in the wage rate can lead to either an increase or a decrease in labor supply.

### 2.11.3 Homothetic Utility Functions

A function  $f: R^n \rightarrow R$  is **homogeneous of degree 1** if  $f(t\mathbf{x}) = tf(\mathbf{x})$  for all  $t > 0$ . A function  $f(\mathbf{x})$  is **homothetic** if  $f(\mathbf{x}) = g(h(\mathbf{x}))$  where  $g$  is a strictly increasing function and  $h$  is a function which is homogeneous of degree 1.

Economists often find it useful to assume that utility functions are homogeneous or homothetic. In fact, there is little distinction between the two concepts in utility theory. A homothetic function is simply a monotonic transformation of a homogeneous function, but utility functions are only defined up to a monotonic transformation. Thus assuming that preferences can be represented by a homothetic function is equivalent to assuming that they can be represented by a function that is homogeneous of degree 1. If a consumer has preferences that can be represented by a homothetic utility function, economists say that the consumer has **homothetic preferences**.

It can easily be shown that if the utility function is homogeneous of degree 1, then the expenditure function can be written as  $e(\mathbf{p}, u) = e(\mathbf{p})u$ . This in turn implies that the indirect utility function can be written as  $v(\mathbf{p}, m) = v(\mathbf{p})m$ . Roy's identity then implies that the demand functions take the form  $x_i(\mathbf{p}, m) = x_i(\mathbf{p})m$  — i.e., they are linear functions of income. The fact that the “income effects” take this special form is often useful in demand analysis, as we will see below.

### 2.11.4 Aggregating Across Goods

In many circumstances it is reasonable to model consumer choice by certain “partial” maximization problems. For example, we may want to model the consumer's choice of “meat” without distinguishing how much is beef, pork, lamb, etc. In most empirical work, some kind of aggregation of this sort is necessary.

In order to describe some useful results concerning this kind of separability of consumption decisions, we will have to introduce some new notation. Let us think of partitioning the consumption bundle into two “subbundles” so that the consumption bundle takes the form  $(\mathbf{x}, \mathbf{z})$ . For example,  $\mathbf{x}$  could be the vector of consumptions of different kinds of



meat, and  $\mathbf{z}$  could be the vector of consumption of all other goods.

We partition the price vector analogously into  $(\mathbf{p}, \mathbf{q})$ . Here  $\mathbf{p}$  is the price vector for the different kinds of meat, and  $\mathbf{q}$  is the price vector for the other goods. With this notation the standard utility maximization problem can be written as

$$\begin{aligned} \max_{\mathbf{x}, \mathbf{z}} u(\mathbf{x}, \mathbf{z}) \\ \text{such that } \mathbf{p}\mathbf{x} + \mathbf{q}\mathbf{z} = m. \end{aligned} \quad (2.24)$$

The problem of interest is under what conditions we can study the demand problem for the  $x$ -goods, say, as a group, without worrying about how demand is divided among the various components of the  $x$ -goods.

One way to formulate this problem mathematically is as follows. We would like to be able to construct some scalar **quantity index**,  $X$ , and some scalar price index,  $P$ , that are functions of the vector of quantities and the vector of prices:

$$\begin{aligned} P &= f(\mathbf{p}) \\ X &= g(\mathbf{x}). \end{aligned} \quad (2.25)$$

In this expression  $P$  is supposed to be some kind of “price index” which gives the “average price” of the goods, while  $X$  is supposed to be a quantity index that gives the average “amount” of meat consumed. Our hope is that we can find a way to construct these price and quantity indices so that they behave like ordinary prices and quantities.

That is, we hope to find a new utility function  $U(X, \mathbf{z})$ , which depends only on the quantity index of  $x$ -consumption, that will give us the same answer as if we solved the entire maximization problem in (2.24). More formally, consider the problem

$$\begin{aligned} \max_{\mathbf{X}, \mathbf{z}} U(X, \mathbf{z}) \\ \text{such that } \mathbf{P}\mathbf{X} + \mathbf{q}\mathbf{z} = m. \end{aligned}$$

The demand function for the quantity index  $X$  will be some function  $X(P, \mathbf{q}, m)$ . We want to know when it will be the case that

$$X(P, \mathbf{q}, m) \equiv X(f(p), \mathbf{q}, m) = g(x(p, \mathbf{q}, m)).$$

This requires that we get to the same value of  $X$  via two different routes:

- 1) first aggregate prices using  $P = f(\mathbf{p})$  and then maximize  $U(X, \mathbf{z})$  subject to the budget constraint  $PX + \mathbf{qz} = m$ .
- 2) first maximize  $u(\mathbf{x}, \mathbf{z})$  subject to  $\mathbf{px} + \mathbf{qz} = m$  and then aggregate quantities to get  $X = g(\mathbf{x})$ .

There are two situations under which this kind of aggregation is possible. The first situation, which imposes constraints on the price movements, is known as **Hicksian separability**. The second, which imposes constraints on the structure of preferences, is known as **functional separability**.

#### Hicksian separability

Suppose that the price vector  $\mathbf{p}$  is always proportional to some fixed base price vector  $\mathbf{p}^0$  so that  $\mathbf{p} = t\mathbf{p}^0$  for some scalar  $t$ . If the  $x$ -goods are various kinds of meat, this condition requires that the relative prices of the various kinds of meat remain constant — they all increase and decrease in the same proportion.

Following the general framework described above, let us define the price and quantity indices for the  $x$ -goods by

$$\begin{aligned} P &= t \\ X &= \mathbf{p}^0 \mathbf{x}. \end{aligned}$$

We define the indirect utility function associated with these indices as

$$\begin{aligned} V(P, \mathbf{q}, m) &= \max_{\mathbf{x}, \mathbf{z}} u(\mathbf{x}, \mathbf{z}) \\ \text{such that } P\mathbf{p}^0 \mathbf{x} + \mathbf{qz} &= m. \end{aligned}$$

It is straightforward to check that this indirect utility function has all the usual properties: it is quasiconvex, homogeneous in price and income, etc. In particular, a straightforward application of the envelope theorem shows that we can recover the demand function for the  $x$ -good by Roy's identity:

$$X(P, \mathbf{q}, m) = -\frac{\partial V(P, \mathbf{q}, m)/\partial P}{\partial V(P, \mathbf{q}, m)/\partial m} = \mathbf{p}^0 \mathbf{x}(\mathbf{p}, \mathbf{q}, m).$$

This calculation shows that  $X(P, \mathbf{q}, m)$  is an appropriate quantity index for the  $x$ -goods consumption: we get the same result if we first aggregate prices and then maximize  $U(X, \mathbf{z})$  as we get if we maximize  $u(\mathbf{x}, \mathbf{z})$  and then aggregate quantities.

We can solve for the direct utility function that is dual to  $V(P, \mathbf{q}, m)$  by the usual calculation:

$$U(X, \mathbf{z}) = \min_{P, \mathbf{q}} V(P, \mathbf{q}, m) \\ \text{such that } PX + \mathbf{q}\mathbf{z} = m.$$

By construction this direct utility function has the property that

$$V(P, \mathbf{q}, m) = \max_{\mathbf{X}, \mathbf{z}} U(X, \mathbf{z}) \\ \text{such that } PX + \mathbf{q}\mathbf{z} = m.$$

Hence, the price and quantity indices constructed this way behave just like ordinary prices and quantities.

### The two-good model

One common application of Hicksian aggregation is when we are studying the demand for a single good. In this case, think of the  $z$ -goods as being a single good,  $z$ , and the  $x$ -goods as “all other goods.” The actual maximization problem is then

$$\max_{\mathbf{x}, z} u(\mathbf{x}, z) \\ \text{such that } \mathbf{p}\mathbf{x} + qz = m.$$

Suppose that the relative prices of the  $x$ -goods remains constant, so that  $\mathbf{p} = P\mathbf{p}^0$ . That is, the vector of prices  $\mathbf{p}$  is some base price vector  $\mathbf{p}^0$  times some price index  $P$ . Then Hicksian aggregation says that we can write the demand function for the  $z$ -good as

$$z = z(P, q, m).$$

Since this demand function is homogeneous of degree zero, with some abuse of notation, we can also write

$$z = z(q/P, m/P).$$

This says that the demand for the  $z$ -good depends on the relative price of the  $z$ -good to “all other goods” and income, divided by the price of “all other goods.” In practice, the price index for all other goods is usually taken to be some standard consumer price index. The demand for the  $z$ -good becomes a function of only two variables: the price of the  $z$ -good relative to the CPI and income relative to the CPI.

### Functional separability

The second case in which we can decompose the consumer's consumption decision is known as the case of functional separability. Let us suppose that the underlying preference ordering has the property that

$$(\mathbf{x}, \mathbf{z}) \succ (\mathbf{x}', \mathbf{z}) \text{ if and only if } (\mathbf{x}, \mathbf{z}') \succ (\mathbf{x}', \mathbf{z}')$$

for all consumption bundles  $\mathbf{x}, \mathbf{x}', \mathbf{z}$  and  $\mathbf{z}'$ . This condition says that if  $\mathbf{x}$  is preferred to  $\mathbf{x}'$  for some choices of the other goods, then  $\mathbf{x}$  is preferred to  $\mathbf{x}'$  for all choices of the other goods. Or, even more succinctly, the preferences over the  $x$ -goods are independent of the  $z$ -goods.

If this “independence” property is satisfied and the preferences are locally nonsatiated, then it can be shown that the utility function for  $\mathbf{x}$  and  $\mathbf{z}$  can be written in the form  $u(\mathbf{x}, \mathbf{z}) = U(v(\mathbf{x}), \mathbf{z})$ , where  $U(v, \mathbf{z})$  is an increasing function of  $v$ . That is, the overall utility from  $\mathbf{x}$  and  $\mathbf{z}$  can be written as a function of the **subutility** of  $\mathbf{x}$ ,  $v(\mathbf{x})$ , and the level of consumption of the  $z$ -goods.

If the utility function can be written in this form, we will say that the utility function is **weakly separable**. What does separability imply about the structure of the utility maximization problem? As usual, we will write the demand function for the goods as  $\mathbf{x}(\mathbf{p}, \mathbf{q}, m)$  and  $\mathbf{z}(\mathbf{p}, \mathbf{q}, m)$ . Let  $m_x = \mathbf{p}\mathbf{x}(\mathbf{p}, \mathbf{q}, m)$  be the optimal expenditure on the  $x$ -goods.

It turns out that if the overall utility function is weakly separable, the optimal choice of the  $x$ -goods can be found by solving the following subutility maximization problem:

$$\begin{aligned} \max v(\mathbf{x}) \\ \text{such that } \mathbf{p}\mathbf{x} &= m_x. \end{aligned} \tag{2.26}$$

This means that if we know the expenditure on the  $x$ -goods,  $m_x = \mathbf{p}\mathbf{x}(\mathbf{p}, \mathbf{q}, m)$ , we can solve the subutility maximization problem to determine the optimal choice of the  $x$ -goods. In other words, the demand for the  $x$ -goods is only a function of the prices of the  $x$ -goods and the expenditure on the  $x$ -goods  $m_x$ . The prices of the *other* goods are only relevant insofar as they determine the expenditure on the  $x$ -goods.

The proof of this is straightforward. Assume that  $\mathbf{x}(\mathbf{p}, \mathbf{q}, m)$  does not solve the above problem. Instead, let  $\mathbf{x}'$  be another value of  $\mathbf{x}$  that satisfies the budget constraint and

yields strictly greater subutility. Then the bundle  $(\mathbf{x}', \mathbf{z})$  would give higher overall utility than  $(\mathbf{x}(\mathbf{p}, \mathbf{q}, m), \mathbf{z}(\mathbf{p}, \mathbf{q}, m))$ , which contradicts the definition of the demand function.

The demand functions  $\mathbf{x}(\mathbf{p}, m_x)$  are sometimes known as conditional **demand functions** since they give demand for the  $x$ -goods conditional on the level of expenditure on these goods. Thus, for example, we may consider the demand for beef as a function of the prices of beef, pork, and lamb and the total expenditure on meat. Let  $e(\mathbf{p}, v)$  be the expenditure function for the subutility maximization problem given in (2.26). This tells us how much expenditure on the  $x$ -goods is necessary at prices  $\mathbf{p}$  to achieve the subutility  $v$ . It is not hard to see that we can write the overall maximization problem of the consumer as

$$\begin{aligned} & \max_{\mathbf{v}, \mathbf{z}} U(v, \mathbf{z}) \\ & \text{such that } e(\mathbf{p}, v) + \mathbf{qz} = m \end{aligned}$$

This is almost in the form we want:  $v$  is a suitable quantity index for the  $x$ -goods, but the price index for the  $x$ -goods isn't quite right. We want  $P$  times  $X$ , but we have some nonlinear function of  $\mathbf{p}$  and  $X = v$ .

In order to have a budget constraint that is linear in quantity index, we need to assume that subutility function has a special structure. For example. Suppose that the subutility function is homothetic. Then we can write  $e(\mathbf{p}, v)$  as  $e(\mathbf{p})v$ . Hence, we can choose our quantity index to be  $X = v(\mathbf{x})$ , our price index to be  $P = e(\mathbf{p})$ , and our utility function to be  $U(X, \mathbf{z})$ . We get the same  $X$  if we solve

$$\begin{aligned} & \max_{X, \mathbf{z}} U(X, \mathbf{z}) \\ & \text{such that } PX + \mathbf{qz} = m \end{aligned}$$

as if we solve

$$\begin{aligned} & \max_{\mathbf{x}, \mathbf{z}} u(v(\mathbf{x}), \mathbf{z}) \\ & \text{such that } \mathbf{px} + \mathbf{qz} = m, \end{aligned}$$

and then aggregate using  $X = v(\mathbf{x})$ .

In this formulation we can think of the consumption decision as taking place in two stages: first the consumer considers how much of the composite commodity (e.g., meat)

to consume as a function of a price index of meat by solving the overall maximization problem; then the consumer considers how much beef to consume given the prices of the various sorts of meat and the total expenditure on meat, which is the solution to the subutility maximization problem. Such a two-stage budgeting process is very convenient in applied demand analysis.

### 2.11.5 Aggregating Across Consumers

We have studied the properties of a consumer's demand function,  $\mathbf{x}(\mathbf{p}, m)$ . Now let us consider some collection of  $i = 1, \dots, n$  consumers, each of whom has a demand function for some  $L$  commodities, so that consumer  $i$ 's demand function is a vector  $\mathbf{x}_i(\mathbf{p}, m_i) = (x_i^1(\mathbf{p}, m_i), \dots, x_i^L(\mathbf{p}, m_i))$  for  $i = 1, \dots, n$ . Note that we have changed our notation slightly: goods are now indicated by superscripts while consumers are indicated by subscripts. The aggregate demand function is defined by  $\mathbf{X}(\mathbf{p}, m_1, \dots, m_n) = \sum_{i=1}^n \mathbf{x}_i(\mathbf{p}, m_i)$ . The aggregate demand for good  $l$  is denoted by  $X^l(\mathbf{p}, \mathbf{m})$  where  $\mathbf{m}$  denotes the vector of incomes  $(m_1, \dots, m_n)$ .

The aggregate demand function inherits certain properties of the individual demand functions. For example, if the individual demand functions are continuous, the aggregate demand function will certainly be continuous. Continuity of the individual demand functions is a sufficient but not necessary condition for continuity of the aggregate demand functions.

What other properties does the aggregate demand function inherit from the individual demands? Is there an aggregate version of Slutsky's equation or of the Strong Axiom of Revealed Preference? Unfortunately, the answer to these questions is no. In fact the aggregate demand function will in general possess no interesting properties other than homogeneity and continuity. Hence, the theory of the consumer places no restrictions on aggregate behavior in general. However, in certain cases it may happen that the aggregate behavior may look as though it were generated by a single "representative" consumer. Below, we consider a circumstance where this may happen.

Suppose that all individual consumers' indirect utility functions take the **Gorman form**:

$$v_i(\mathbf{p}, m_i) = a_i(\mathbf{p}) + b(\mathbf{p})m_i.$$

Note that the  $a_i(\mathbf{p})$  term can differ from consumer to consumer, but the  $b(\mathbf{p})$  term is assumed to be identical for all consumers. By Roy's identity the demand function for good  $j$  of consumer  $i$  will then take the form

$$x_i^j(\mathbf{p}, m_i) = a_i^j(\mathbf{p}) + \beta^j(\mathbf{p})m_i. \quad (2.27)$$

where,

$$\begin{aligned} a_i^j(\mathbf{p}) &= -\frac{\frac{\partial a_i(\mathbf{p})}{\partial p_j}}{b(\mathbf{p})} \\ \beta^j(\mathbf{p}) &= -\frac{\frac{\partial b(\mathbf{p})}{\partial p_j}}{b(\mathbf{p})}. \end{aligned}$$

Note that the marginal propensity to consume good  $j$ ,  $\partial x_i^j(\mathbf{p}, m_i)/\partial m_i$ , is independent of the level of income of any consumer and also constant across consumers since  $b(\mathbf{p})$  is constant across consumers. The aggregate demand for good  $j$  will then take the form

$$X^j(\mathbf{p}, m^1, \dots, m^n) = - \left[ \sum_{i=1}^n \frac{\frac{\partial a_i}{\partial p_j}}{b(\mathbf{p})} + \frac{\frac{\partial b(p)}{\partial p_j}}{b(\mathbf{p})} \sum_{i=1}^n m_i \right].$$

This demand function can in fact be generated by a representative consumer. His representative indirect utility function is given by

$$V(\mathbf{p}, M) = \sum_{i=1}^n a_i(\mathbf{p}) + b(\mathbf{p})M = A(\mathbf{p}) + B(\mathbf{p})M,$$

where  $M = \sum_{i=1}^n m_i$ .

The proof is simply to apply Roy's identity to this indirect utility function and to note that it yields the demand function given in equation (2.27). In fact it can be shown that the Gorman form is the most general form of the indirect utility function that will allow for aggregation in the sense of the representative consumer model. Hence, the Gorman form is not only *sufficient* for the representative consumer model to hold, but it is also *necessary*.

Although a complete proof of this fact is rather detailed, the following argument is reasonably convincing. Suppose, for the sake of simplicity, that there are only two consumers. Then by hypothesis the aggregate demand for good  $j$  can be written as

$$X^j(\mathbf{p}, m_1 + m_2) \equiv x_1^j(\mathbf{p}, m_1) + x_2^j(\mathbf{p}, m_2).$$

If we first differentiate with respect to  $m_1$  and then with respect to  $m_2$ , we find the following identities

$$\frac{\partial^2 X^j(\mathbf{p}, M)}{\partial M} \equiv \frac{\partial x_1^j(\mathbf{p}, m_1)}{\partial m_1} \equiv \frac{\partial x_2^j(\mathbf{p}, m_2)}{\partial m_2}.$$

Hence, the marginal propensity to consume good  $j$  must be the same for all consumers. If we differentiate this expression once more with respect to  $m_1$ , we find that

$$\frac{\partial^2 X^j(\mathbf{p}, M)}{\partial M^2} \equiv \frac{\partial^2 x_1^j(\mathbf{p}, m_1)}{\partial m_1^2} \equiv 0.$$

Thus, consumer 1's demand for good  $j$  – and, therefore, consumer 2's demand – is affine in income. Hence, the demand functions for good  $j$  take the form  $x_i^j(\mathbf{p}, m_i) = a_i^j(\mathbf{p}) + \beta^j(\mathbf{p})m_i$ . If this is true for all goods, the indirect utility function for each consumer must have the Gorman form.

One special case of a utility function having the Gorman form is a utility function that is homothetic. In this case the indirect utility function has the form  $v(\mathbf{p}, m) = v(\mathbf{p})m$ , which is clearly of the Gorman form. Another special case is that of a quasi-linear utility function. A utility function  $U$  is said to be quasi-linear if it has the following functional form:

$$U(x_0, x_1, \dots, x_L) = x_0 + u(x_1, \dots, x_L)$$

In this case  $v(\mathbf{p}, m) = v(\mathbf{p}) + m$ , which obviously has the Gorman form. Many of the properties possessed by homothetic and/or quasi-linear utility functions are also possessed by the Gorman form.

The class of quasi-linear utility function is a very important class of utility functions which plays an important role in many economics fields such as those of information economics, mechanism design theory, property rights theory due to its important property of no income effect when income changes.

## Reference

- Afriat, S. (1967). The construction of a utility function from expenditure data. *International Economic Review*, 8, 67-77.
- Debreu, G., *Theory of Value*, (Wiley, New York), 1959.



- Debreu, G. (1964). Continuity properties of Paretian utility. *International Economic Review*, 5, 285 -293.
- Gorman, T. (1953). Community preference fields. *Econometrica*, 21, 63-80.
- Hurwicz, L. & Uzawa, H. (1971). On the integrability of demand functions. In J. Chipman, L. Hurwicz, M. Richter & H. Sonnenschein (Eds.), *Preferences, Utility, and Demand*. (New York: Harcourt, Brace, Jovanovich).
- Jehle, G. A., and P. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, 1998, Chapters 3-4.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 7.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995, Chapters 1-4.
- Roy, R. (1942). *De l'utilité* (Paris: Hermann).
- Roy, R. (1947). La distribution de revenu entre les divers biens. *Econometrica*, 15, 205-225.
- Rubinstein Ariel, *Lecture Notes in Microeconomics (modeling the economic agent)*, Princeton University Press, 2005.
- Samuelson, P. (1948). Consumption theory in terms of revealed preference. *Econometrica*, 15, 243-253.
- Sonnenschein, H., *Demand Theory without Transitive Preferences, with Application to the Theory of Competitive Equilibrium*, Preferences, Utility, and Demand, Edited. by J. S. Chipman, L. Hurwicz, M. K. Richter, and H. Sonnenschein, Harcourt Brace Jovanovich, New York, New York, 1971.
- Takayama, A. *Mathematical Economics*, the second edition, Cambridge: Cambridge University Press, 1985, Chapters 1-3.
- Tian, G., *Market Economics for Masses*, (with Fan Zhang), in A Series of Market Economics, Vol. 1, Ed. by G. Tian, Shanghai People's Publishing House and Hong Kong's Intelligent Book Ltd, 1993 (in Chinese), Chapters 1-2.

- Tian, G. “Generalizations of the FKKM Theorem and Ky-Fan Minimax Inequality, with Applications to Maximal Elements, Price Equilibrium, and Complementarity,” *Journal of Mathematical Analysis and Applications*, 170 (1992), pp. 457-471.
- Tian, G., “Generalized KKM Theorem and Minimax Inequalities and Their Applications,” *Journal of Optimization Theory and Applications*, 83 (1994), 375-389.
- Tian, G. “Necessary and Sufficient Conditions for Maximization of a Class of Preference Relations,” *Review of Economic Studies*, 60 (1993), 949-958.
- “Generalizations of the FKKM Theorem and Ky-Fan Minimax Inequality, with Applications to Maximal Elements, Price Equilibrium, and Complementarity,” *Journal of Mathematical Analysis and Applications*, 170 (1992), pp. 457-471.
- Tian, G. and J. Zhou, “Transfer Method for Characterizing the Existence of Maximal Elements of Binary Relations on Compact or Noncompact Sets,” *SIAM Journal on Optimization*, 2 (1992), pp. 360-375.
- Tian, G. and J. Zhou, “Transfer Continuities, Generalizations of the Weierstrass Theorem and Maximum Theorem—A Full Characterization,” *Journal of Mathematical Economics*, 24 (1995), 281-303.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 7-9.
- Wold, H. (1943). A synthesis of pure demand analysis, I-III. *Skandinavisk Aktuarietidskrift*, 26,27.

# Chapter 3

## Production Theory

### 3.1 Introduction

Economic activity not only involves consumption but also production and trade. Production should be interpreted very broadly, however, to include production of both physical goods – such as rice or automobiles–and services–such as medical care or financial services.

A firm can be characterized by many factors and aspects such as sectors, production scale, ownerships, organization structures, etc. But which are most important features for us to study producer's behavior in making choices? To grasp the most important features in studying producer behavior and choices in modern producer theory, it is assumed that the key characteristic of a firm is production set. Producer's characteristic together with the behavior assumption are building blocks in any model of producer theory. The production set represents the set of all technologically feasible production plans. The behavior assumption expresses the guiding principle the producer uses to make choices. It is generally assumed that the producer seeks to identify and select a production that is most profitable.

We will first present a general framework of production technology. By itself, the framework does not describe how production choices are made. It only specifies basic characteristic of a firm which defines what choices can be made; it does not specify what choices should be made. We then will discuss what choices should be made based on the behavior assumptions on firms. A basic behavior assumption on producers is profit maximization. After that, we will describe production possibilities in physical terms,

which is recast into economic terms – using cost functions.

## 3.2 Production Technology

Production is the process of transforming inputs to outputs. Typically, inputs consist of labor, capital equipment, raw materials, and intermediate goods purchased from other firms. Outputs consists of finished products or service, or intermediate goods to be sold to other firms. Often alternative methods are available for producing the same output, using different combinations of inputs. A firm produces outputs from various combinations of inputs. In order to study firm choices we need a convenient way to summarize the production possibilities of the firm, i.e., which combinations of inputs and outputs are **technologically feasible**.

### 3.2.1 Measurement of Inputs and Outputs

It is usually most satisfactory to think of the inputs and outputs as being measured in terms of flows: a certain amount of inputs per the period are used to produce a certain amount of outputs per unit the period at some location. It is a good idea to explicitly include the time and location dimensions in a specification of inputs and outputs. The level of detail that we will use in specifying inputs and outputs will depend on the problem at hand, but we should remain aware of the fact that a particular input or output good can be specified in arbitrarily fine detail. However, when discussing technological choices in the abstract, as we do in this chapter, it is common to omit the time and location dimensions.

### 3.2.2 Specification of Technology

The fundamental reality firms must contend with in this process is technological feasibility. The state of technology determines and restricts what is possible in combining inputs to produce outputs, and there are several way we can represent this constraint. The most general way is to think of the firm as having a production possibility set.

Suppose the firm has  $L$  possible goods to serve as inputs and/or outputs. If a firm uses  $y_j^i$  units of a good  $j$  as an input and produces  $y_j^o$  of the good as an output, then the

**net output** of good  $j$  is given by  $y_j = y_j^o - y_j^i$ .

A **production plan** is simply a list of net outputs of various goods. We can represent a production plan by a vector  $y$  in  $R^L$  where  $y_j$  is negative if the  $j^{th}$  good serves as a net input and positive if the  $j^{th}$  good serves as a net output. The set of all technologically feasible production plans is called the firm's **production possibilities set** and will be denoted by  $Y$ , a subset of  $R^L$ . The set  $Y$  is supposed to describe all patterns of inputs and outputs that are technologically feasible. It gives us a complete description of the technological possibilities facing the firm.

When we study the behavior of a firm in certain economic environments, we may want to distinguish between production plans that are “immediately feasible” and those that are “eventually” feasible. We will generally assume that such restrictions can be described by some vector  $z$  in  $R^L$ . The **restricted** or **short-run production possibilities set** will be denoted by  $Y(z)$ ; this consists of all feasible net output bundles consistent with the constraint level  $z$ . The following are some examples of such restrictions.

#### EXAMPLE: INPUT REQUIREMENT SET

Suppose a firm produces only one output. In this case we write the net output bundle as  $(y, -\mathbf{x})$  where  $\mathbf{x}$  is a vector of inputs that can produce  $y$  units of output. We can then define a special case of a restricted production possibilities set, the **input requirement set**:

$$V(y) = \{\mathbf{x} \text{ in } R_+^L: (y, -\mathbf{x}) \text{ is in } Y\}$$

The input requirement set is the set of all input bundles that produce at least  $y$  units of output.

Note that the input requirement set, as defined here, measures inputs as positive numbers rather than negative numbers as used in the production possibilities set.

#### EXAMPLE: ISOQUANT

In the case above we can also define an **isoquant**:

$$Q(y) = \{\mathbf{x} \text{ in } R_+^n: \mathbf{x} \text{ is in } V(y) \text{ and } \mathbf{x} \text{ is not in } V(y') \text{ for } y' > y\}.$$

The isoquant gives all input bundles that produce exactly  $y$  units of output.

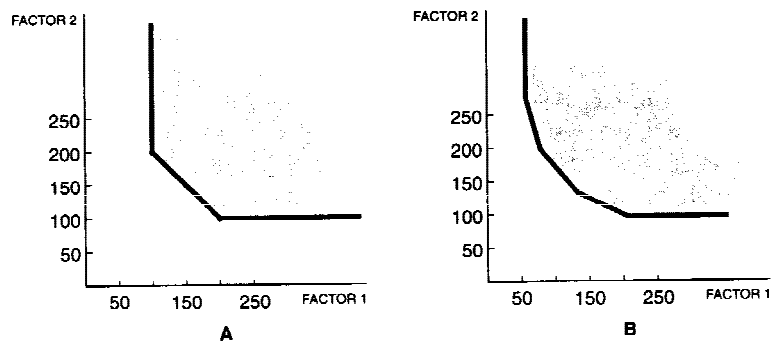


Figure 3.1: Convex input requirement sets.

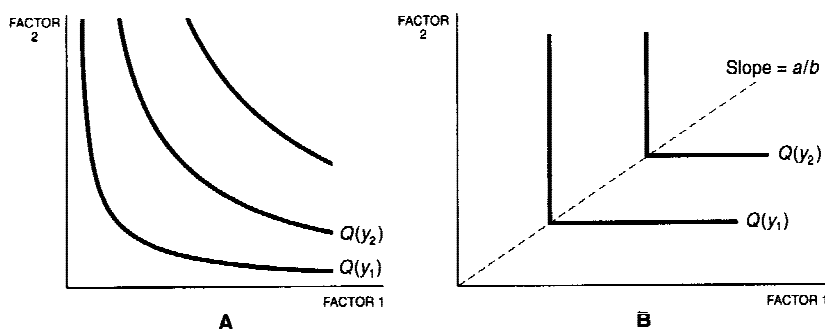


Figure 3.2: Panel A depicts the general shape of a isoquant curve, and panel B depicts the general shape of a perfect complement (Leontief) technology.

#### EXAMPLE: SHORT-RUN PRODUCTION POSSIBILITIES SET

Suppose a firm produces some output from labor and some kind of machine which we will refer to as “capital.” Production plans then look like  $(y, -l, -k)$  where  $y$  is the level of output,  $l$  the amount of labor input, and  $k$  the amount of capital input. We imagine that labor can be varied immediately but that capital is fixed at the level  $\bar{k}$  in the short run. Then

$$Y(\bar{k}) = \{(y, -l, -k) \text{ in } Y: k = \bar{k}\}$$

is a **short-run production possibilities set**.

#### EXAMPLE: PRODUCTION FUNCTION

If the firm has only one output, we can define the **production function**:  $f(\mathbf{x}) = \{y \text{ in } R: y \text{ is the maximum output associated with } -\mathbf{x} \text{ in } Y\}$ .

#### EXAMPLE: TRANSFORMATION FUNCTION

A production plan  $y$  in  $Y$  is (**technologically**) efficient if there is no  $\mathbf{y}'$  in  $Y$  such that  $\mathbf{y}' \geq \mathbf{y}$  and  $\mathbf{y}' \neq \mathbf{y}$ ; that is, a production plan is efficient if there is no way to produce more output with the same inputs or to produce the same output with less inputs. (Note carefully how the sign convention on inputs works here.) We often assume that we can describe the set of technologically efficient production plans by a **transformation function**  $T: \rightarrow R$  where  $T(\mathbf{y}) = 0$  if and only if  $\mathbf{y}$  is efficient. Just as a production function picks out the maximum *scalar* output as a function of the inputs, the transformation function picks out the maximal *vectors* of net outputs.

#### EXAMPLE: COBB-DOUGLAS TECHNOLOGY

Let  $\alpha$  be a parameter such that  $0 < \alpha < 1$ . Then the **Cobb-Douglas technology** is defined in the following manner.

$$\begin{aligned} Y &= \{(y, -x_1, -x_2) \text{ in } R^3: y \leq x_1^\alpha x_2^{1-\alpha}\} \\ V(y) &= \{(x_1, x_2) \text{ in } R_+^2: y \leq x_1^\alpha x_2^{1-\alpha}\} \\ Q(y) &= \{(x_1, x_2) \text{ in } R_+^2: y = x_1^\alpha x_2^{1-\alpha}\} \\ Y(z) &= \{(y, -x_1, -x_2) \text{ in } R^3: y \leq x_1^\alpha x_2^{1-\alpha}, x_2 = z\} \\ T(y, x_1 x_2) &= y - x_1^\alpha x_2^{1-\alpha} \\ f(x_1, x_2) &= x_1^\alpha x_2^{1-\alpha}. \end{aligned}$$

#### EXAMPLE: LEONTIEF TECHNOLOGY

Let  $a > 0$  and  $b > 0$  be parameters. Then the **Leontief technology** is defined in the following manner.

$$\begin{aligned} Y &= \{(y, -x_1, -x_2) \text{ in } R^3: y \leq \min(ax_1, bx_2)\} \\ V(y) &= \{(x_1, x_2) \text{ in } R_+^2: y \leq \min(ax_1, bx_2)\} \\ Q(y) &= \{(x_1, x_2) \text{ in } R_+^2: y = \min(ax_1, bx_2)\} \\ T(y, x_1, x_2) &= y - \min(ax_1, bx_2) \\ f(x_1, x_2) &= \min(ax_1, bx_2). \end{aligned}$$

### 3.2.3 Common Properties of Production Sets

Although the production possibility sets of different processes can differ widely in structure, many technologies share certain general properties. If it can be assumed that these properties are satisfied, special theoretical results can be derived. Some important properties are defined below:

**POSSIBILITY OF INACTION:**  $0 \in Y$ .

Possibility of inaction means that no action on production is a possible production plan.

**Closeness:**  $Y$  is closed.

The possibility set  $Y$  is closed means that, whenever a sequence of production plans  $y_i$ ,  $i = 1, 2, \dots$ , are in  $Y$  and  $y_i \rightarrow y$ , then the limit production plan  $y$  is also in  $Y$ . It guarantees that points on the boundary of  $Y$  are feasible. Note that  $Y$  is closed implies that the input requirement set  $V(y)$  is a closed set for all  $y \geq 0$ .

**FREE DISPOSAL OR MONOTONICITY:** If  $y \in Y$  implies that  $y' \in Y$  for all  $y' \leq y$ , then the set  $Y$  is said to satisfy the free disposal or monotonicity property.

Free disposal implies that commodities (either inputs or outputs) can be thrown away. This property means that if  $y \in Y$ , then  $Y$  includes all vectors in the negative orthant translated to  $y$ , i.e. there are only inputs, but no outputs.

A weaker requirement is that we only assume that the input requirement is monotonic: If  $\mathbf{x}$  is in  $V(y)$  and  $\mathbf{x}' \geq \mathbf{x}$ , then  $\mathbf{x}'$  is in  $V(y)$ . Monotonicity of  $V(y)$  means that, if  $\mathbf{x}$  is a feasible way to produce  $y$  units of output and  $\mathbf{x}'$  is an input vector with at least as much of each input, then  $\mathbf{x}'$  should be a feasible way to produce  $y$ .

**IRREVERSIBILITY:**  $Y \cap \{-Y\} = \{0\}$ .

Irreversibility means a production plan is not reversible unless it is a non-action plan.

**CONVEXITY:**  $Y$  is convex if whenever  $y$  and  $y'$  are in  $Y$ , the weighted average  $ty + (1 - t)y'$  is also in  $Y$  for any  $t$  with  $0 \leq t \leq 1$ .



Convexity of  $Y$  means that, if all goods are divisible, it is often reasonable to assume that two production plans  $y$  and  $y'$  can be scaled downward and combined. However, it should be noted that the convexity of the production set is a strong hypothesis. For example, convexity of the production set rules out “start up costs” and other sorts of returns to scale. This will be discussed in greater detail shortly.

**STRICT CONVEXITY:**  $y$  is strictly convex if  $y \in Y$  and  $y' \in Y$ , then  $ty + (1 - t)y' \in \text{int}Y$  for all  $0 < t < 1$ , where  $\text{int}Y$  denotes the interior points of  $Y$ .

As we will show, the strict convexity of  $Y$  can guarantee the profit maximizing production plan is unique provided it exists.

A weak and more reasonable requirement is to assume that  $V(y)$  is a convex set for all outputs  $y_o$ :

**CONVEXITY OF INPUT REQUIREMENT SET:** *If  $\mathbf{x}$  and  $\mathbf{x}'$  are in  $V(y)$ , then  $t\mathbf{x} + (1 - t)\mathbf{x}'$  is in  $V(y)$  for all  $0 \leq t \leq 1$ . That is,  $V(y)$  is a **convex set**.*

Convexity of  $V(y)$  means that, if  $x$  and  $x'$  both can produce  $y$  units of output, then any weighted average  $tx + (1 - t)x'$  can also produce  $y$  units of output. We describe a few of the relationships between the convexity of  $V(y)$ , the curvature of the production function, and the convexity of  $Y$ . We first have

**Proposition 3.2.1 (Convex Production Set Implies Convex Input Requirement Set)**

*If the production set  $Y$  is a convex set, then the associated input requirement set,  $V(y)$ , is a convex set.*

PROOF. If  $Y$  is a convex set then it follows that for any  $\mathbf{x}$  and  $\mathbf{x}'$  such that  $(y, -\mathbf{x})$  and  $(y, -\mathbf{x}')$  are in  $Y$  for  $0 \leq t \leq 1$ , we must have  $(ty + (1 - t)y, t\mathbf{x} - (1 - t)\mathbf{x}')$  in  $Y$ . This is simply requiring that  $(y, (t\mathbf{x} + (1 - t)\mathbf{x}'))$  is in  $Y$ . It follows that if  $x$  and  $x'$  are in  $V(y)$ ,  $t\mathbf{x} + (1 - t)\mathbf{x}'$  is in  $V(y)$  which shows that  $V(y)$  is convex.

**Proposition 3.2.2**  *$V(y)$  is a convex set if and only if the production function  $f(\mathbf{x})$  is a quasi-concave function.*

PROOF  $V(y) = \{\mathbf{x}: f(\mathbf{x}) \geq y\}$ , which is just the upper contour set of  $f(\mathbf{x})$ . But a function is quasi-concave if and only if it has a convex upper contour set.

### 3.2.4 Returns to Scale

Suppose that we are using some vector of inputs  $\mathbf{x}$  to produce some output  $y$  and we decide to scale all inputs up or down by some amount  $t \geq 0$ . What will happen to the level of output? The notions of returns to scale can be used to answer this question. Returns to scale refer to how output responds when all inputs are varied in the same proportion so that they consider long run production processes. There are three possibilities: technology exhibits (1) **constant returns to scale**; (2) **decreasing returns to scale**, and (3) **increasing returns to scale**. Formally, we have

(GLOBAL) RETURNS TO SCALE. *A production function  $f(x)$  is said to exhibit :*

- (1) **constant returns to scale** if  $f(t\mathbf{x}) = tf(\mathbf{x})$  for all  $t \geq 0$ ;
- (2) **decreasing returns to scale** if  $f(t\mathbf{x}) < tf(\mathbf{x})$  for all  $t > 1$ ;
- (3) **increasing returns to scale** if  $f(t\mathbf{x}) > tf(\mathbf{x})$  for all  $t > 1$ ;

Constant returns to scale (CRS) means that doubling inputs exactly double outputs, which is often a reasonable assumption to make about technologies. Decreasing returns to scale means that doubling inputs are less than doubling outputs. Increasing returns to scale means that doubling inputs are more than doubling outputs.

Note that a technology has constant returns to scale if and only if its production function is homogeneous of degree 1. Constant returns to scale is also equivalent to: the statement  $\mathbf{y}$  in  $Y$  implies  $t\mathbf{y}$  is in  $Y$  for all  $t \geq 0$ ; or equivalent to the statement  $\mathbf{x}$  in  $V(y)$  implies  $t\mathbf{x}$  is in  $V(ty)$  for all  $t > 1$ .

It may be remarked that the various kinds of returns to scale defined above are global in nature. It may well happen that a technology exhibits increasing returns to scale for some values of  $\mathbf{x}$  and decreasing returns to scale for other values. Thus in many circumstances a local measure of returns to scale is useful. To define locally returns to scale, we first define elasticity of scale.

The **elasticity of scale** measures the percent increase in output due to a one percent increase in all inputs— that is, due to an increase in the scale of operations.

Let  $y = f(x)$  be the production function. Let  $t$  be a positive scalar, and (consider the function  $y(t) = f(tx)$ ). If  $t = 1$ , we have the current scale of operation; if  $t > 1$ , we are

scaling all inputs up by  $t$ ; and if  $t < 1$ , we are scaling all inputs down by  $t$ . The elasticity of scale is given by

$$e(\mathbf{x}) = \frac{\frac{dy(t)}{y(t)}}{\frac{dt}{t}},$$

evaluated at  $t = 1$ . Rearranging this expression, we have

$$e(\mathbf{x}) = \left. \frac{dy(t)}{dt} \frac{t}{y} \right|_{t=1} = \left. \frac{df(t\mathbf{x})}{dt} \frac{t}{f(t\mathbf{x})} \right|_{t=1}$$

Note that we must evaluate the expression at  $t = 1$  to calculate the elasticity of scale at the point  $\mathbf{x}$ .

Thus, we have the following the local returns to scale:

(LOCAL) RETURNS TO SCALE. A production function  $f(x)$  is said to exhibit **locally increasing, constant, or decreasing returns to scale** as  $e(\mathbf{x})$  is greater, equal, or less than 1.

### 3.2.5 The Marginal Rate of Technical Substitution

Suppose that technology is summarized by a smooth production function and that we are producing at a particular point  $y^* = f(x_1^*, x_2^*)$ . Suppose that we want to increase a small amount of input 1 and decrease some amount of input 2 so as to maintain a constant level of output. How can we determine this **marginal rate of technical substitution (MRTS)** between these two factors? The way is the same as for deriving the marginal rate of substitution of an indifference curve. Differentiating production function when output keeps constant, we have

$$0 = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2,$$

which can be solved for

$$\frac{dx_2}{dx_1} = -\frac{\partial f/\partial x_1}{\partial f/\partial x_2} \equiv -\frac{MP_{x_1}}{MP_{x_2}}.$$

This gives us an explicit expression for the marginal rate technical substitution, which is the rate of marginal production of  $x_1$  and marginal production of  $x_2$ .

**Example 3.2.1 (MRTS for a Cobb-Douglas Technology)** Given that  $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ , we can take the derivatives to find

$$\begin{aligned}\frac{\partial f(x)}{\partial x_1} &= \alpha x_1^{\alpha-1} x_2^{1-\alpha} = \alpha \left[ \frac{x_2}{x_1} \right]^{1-\alpha} \\ \frac{\partial f(x)}{\partial x_2} &= (1-\alpha) x_1^\alpha x_2^{-\alpha} = (1-\alpha) \left[ \frac{x_1}{x_2} \right]^\alpha.\end{aligned}$$

It follows that

$$\frac{\partial x_2(x_1)}{\partial x_1} = -\frac{\partial f / \partial x_1}{\partial f / \partial x_2} = -\frac{\alpha}{1-\alpha} \frac{x_1}{x_2}.$$

### 3.2.6 The Elasticity of Substitution

The marginal rate of technical substitution measures the slope of an isoquant. The **elasticity of substitution** measures the curvature of an isoquant. More specifically, the elasticity of substitution measures the percentage change in the factor ratio divided by the percentage change in the MRTS, with output being held fixed. If we let  $\Delta(x_2/x_1)$  be the change in the factor ratio and  $\Delta \text{MRTS}$  be the change in the technical rate of substitution, we can express this as

$$\sigma = \frac{\frac{\Delta(x_2/x_1)}{x_2/x_1}}{\frac{\Delta \text{MRTS}}{\text{MRTS}}}.$$

This is a relatively natural measure of curvature: it asks how the ratio of factor inputs changes as the slope of the isoquant changes. If a small change in slope gives us a large change in the factor input ratio, the isoquant is relatively flat which means that the elasticity of substitution is large.

In practice we think of the percent change as being very small and take the limit of this expression as  $\Delta$  goes to zero. Hence, the expression for  $\sigma$  becomes

$$\sigma = \frac{\text{MRTS}}{(x_2/x_1)} \frac{d(x_2/x_1)}{d \text{MRTS}} = \frac{d \ln(x_2/x_1)}{d \ln |\text{MRTS}|}.$$

(The absolute value sign in the denominator is to convert the  $\text{MRTS}$  to a positive number so that the logarithm makes sense.)

**Example 3.2.2 (The Cobb-Douglas Production Function)** We have seen above that

$$\text{MRTS} = -\frac{\alpha}{1-\alpha} \frac{x_2}{x_1},$$

or

$$\frac{x_2}{x_1} = -\frac{1-a}{a}MRTS.$$

It follows that

$$\ln \frac{x_2}{x_1} = \ln \frac{1-a}{a} + \ln |MRTS|.$$

This in turn implies

$$\sigma = \frac{d \ln(x_2/x_1)}{d \ln |MRTS|} = 1.$$

**Example 3.2.3 (The CES Production Function)** The constant elasticity of substitution (CES) production function has the form

$$y = [a_1x_1^\rho + a_2x_2^\rho]^{\frac{1}{\rho}}.$$

It is easy to verify that the CES function exhibits constant returns to scale. It will probably not surprise you to discover that the CES production function has a constant elasticity of substitution. To verify this, note that the marginal rate of technical substitution is given by

$$MRTS = -\left(\frac{x_1}{x_2}\right)^{\rho-1},$$

so that

$$\frac{x_2}{x_1} = |MRTS|^{\frac{1}{1-\rho}}.$$

Taking logs, we see that

$$\ln \frac{x_2}{x_1} = \frac{1}{1-\rho} \ln |MRTS|.$$

Applying the definition of  $\sigma$  using the logarithmic derivative,

$$\sigma = \frac{d \ln x_2/x_1}{d \ln |MRTS|} = \frac{1}{1-\rho}.$$

## 3.3 Profit Maximization

### 3.3.1 Producer Behavior

A basic hypothesis on individual firm behavior in the producer theory is that a firm will always choose a most profitable production plan from the production set. We will derive input demand and output supply functions by considering a model of profit-maximizing behavior coupled with a description of underlying production constraints.

Economic **profit** is defined to be the difference between the revenue a firm receives and the costs that it incurs. It is important to understand that all (explicit and implicit) costs must be included in the calculation of profit. Both revenues and costs of a firm depend on the actions taken by the firm. We can write revenue as a function of the level of operations of some  $n$  actions,  $R(a_1, \dots, a_n)$ , and costs as a function of these same  $n$  activity levels,  $C(a_1, \dots, a_n)$ , where actions can be in term of employment level of inputs or output level of production or prices of outputs if the firm has a market power to set up the prices.

A basic assumption of most economic analysis of firm behavior is that a firm acts so as to maximize its profits; that is, a firm chooses actions  $(a_1, \dots, a_n)$  so as to maximize  $R(a_1, \dots, a_n) - C(a_1, \dots, a_n)$ . The profit maximization problem facing the firm can be then written as

$$\max_{a_1, \dots, a_n} R(a_1, \dots, a_n) - C(a_1, \dots, a_n).$$

The first order conditions for interior optimal actions,  $\mathbf{a}^* = (a_1^*, \dots, a_n^*)$ , is characterized by the conditions

$$\frac{\partial R(\mathbf{a}^*)}{\partial a_i} = \frac{\partial C(\mathbf{a}^*)}{\partial a_i} \quad i = 1, \dots, n.$$

The intuition behind these conditions should be clear: if marginal revenue were greater than marginal cost, it would pay to increase the level of the activity; if marginal revenue were less than marginal cost, it would pay to decrease the level of the activity. In general, revenue is composed of two parts: how much a firm sells of various outputs times the price of each output. Costs are also composed of two parts: how much a firm uses of each input times the price of each input.

The firm's profit maximization problem therefore reduces to the problem of determining what prices it wishes to charge for its outputs or pay for its inputs, and what levels of outputs and inputs it wishes to use. In determining its optimal policy, the firm faces two kinds of constraints: technological constraints that are specified by production sets and market constraints that concern the effect of actions of other agents on the firm. The firms described in the remainder of this chapter are assumed to exhibit the simplest kind of market behavior, namely that of **price-taking** behavior. Each firm will be assumed to take prices as given. Thus, the firm will be concerned only with determining the profit-maximizing levels of outputs and inputs. Such a price-taking firm is often referred to as

a **competitive firm**. We will consider the general case in Chapter 6 – the theory of markets.

### 3.3.2 Producer's Optimal Choice

Let  $\mathbf{p}$  be a vector of prices for inputs and outputs of the firm. The profit maximization problem of the firm can be stated

$$\begin{aligned} \pi(\mathbf{p}) = \max \mathbf{p}\mathbf{y} \\ \text{such that } \mathbf{y} \text{ is in } Y. \end{aligned} \tag{3.1}$$

Note that since outputs are measured as positive numbers and inputs are measured as negative numbers, the objective function for this problem is profits: revenues minus costs. The function  $\pi(\mathbf{p})$ , which gives us the maximum profits as a function of the prices, is called the **profit function** of the firm.

There are several useful variants of the profit function:

Case 1. Short-run maximization problem. In this case, we might define the **short-run profit function**, also known as the **restricted profit function**:

$$\begin{aligned} \pi(\mathbf{p}, \mathbf{z}) = \max \mathbf{p}\mathbf{y} \\ \text{such that } \mathbf{y} \text{ is in } Y(\mathbf{z}). \end{aligned}$$

Case 2. If the firm produces only one output, the profit function can be written as

$$\pi(p, \mathbf{w}) = \max p f(\mathbf{x}) - \mathbf{w}\mathbf{x}$$

where  $p$  is now the (scalar) price of output,  $\mathbf{w}$  is the vector of factor prices, and the inputs are measured by the (nonnegative) vector  $\mathbf{x} = (x_1, \dots, x_n)$ .

The value of  $\mathbf{y}$  that solves the profit problem (3.1) is in general not unique. When there is such a unique production plan, the production plan is called the **net output function** or **net supply function**, the corresponding input part is called the producer's **input demand function** and the corresponding output vector is called the producer's **output supply function**. We will see from the following proposition that strict convexity of production set will ensure the uniqueness of optimal production plan.

**Proposition 3.3.1** *Suppose  $Y$  strictly convex. Then, for each given  $p \in \mathfrak{R}_+^L$ , the profit maximizing production is unique provide it exists.*

Proof: Suppose not. Let  $y$  and  $y'$  be two profit maximizing production plans for  $p \in \mathfrak{R}_+^L$ . Then, we must have  $py = py'$ . Thus, by the strict convexity of  $Y$ , we have  $ty + (1 - t)y' \in \text{int}Y$  for all  $0 < t < 1$ . Therefore, there exists some  $k > 1$  such that

$$k[ty + (1 - t)y'] \in \text{int}Y. \quad (3.2)$$

Then  $k[tpy + (1 - t)py'] = kpy > py$  which contradicts the fact that  $y$  is a profit maximizing production plan.

### 3.3.3 Producer's First-Order Conditions

Profit-maximizing behavior can be characterized by calculus when the technology can be described by a differentiable production function. For example, the first-order conditions for the single output profit maximization problem with interior solution are

$$p \frac{\partial f(\mathbf{x}^*)}{\partial x_i} = w_i \quad i = 1, \dots, n. \quad (3.3)$$

Using vector notation, we can also write these conditions as

$$p\mathbf{D}f(\mathbf{x}^*) = \mathbf{w}.$$

The first-order conditions state that the “marginal value of product of each factor must be equal to its price,” i.e., marginal revenue equals marginal cost at the profit maximizing production plan. This first-order condition can also be exhibited graphically. Consider the production possibilities set depicted in Figure 3.3. In this two-dimensional case, profits are given by  $\Pi = py - wx$ . The level sets of this function for fixed  $p$  and  $w$  are straight lines which can be represented as functions of the form:  $y = \Pi/p + (w/p)x$ . Here the slope of the isoprofit line gives the wage measured in units of output, and the vertical intercept gives us profits measured in units of output. The point of maximal profits the production function must lie below its tangent line at  $x^*$ ; i.e., it must be “locally concave.”

Similar to the arguments in the consumer theory, the calculus conditions derived above make sense only when the choice variables can be varied in an open neighborhood of the



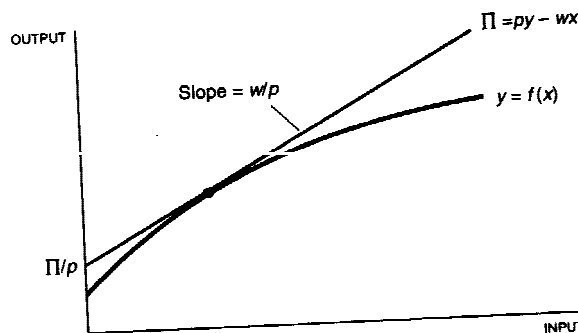


Figure 3.3: Profit maximization when the slope of the isoprofit line equals the slope of the production function.

optimal choice. The relevant first-order conditions that also include **boundary solutions** are given by the Kuhn-Tucker conditions:

$$p \frac{\partial f(\mathbf{x})}{\partial x_i} - w_i \leq 0 \quad \text{with equality if } x_i > 0 \quad (3.4)$$

**Remark 3.3.1** There may exist no profit maximizing production plan when a production technology exhibits constant returns to scale or increasing returns to scale. For example, consider the case where the production function is  $f(x) = x$ . Then for  $p > w$  no profit-maximizing plan will exist. It is clear from this example that the only nontrivial profit-maximizing position for a constant-returns-to-scale firm is the case of  $p = w$  and zero profits. In this case, all production plans are profit-maximizing production plans. If  $(y, \mathbf{x})$  yields maximal profits of zero for some constant returns technology, then  $(ty, t\mathbf{x})$  will also yield zero profits and will therefore also be profit-maximizing.

### 3.3.4 Sufficiency of Producer's First-Order Condition

The second-order condition for profit maximization is that the matrix of second derivatives of the production function must be negative **semi-definite** at the optimal point; that is, the second-order condition requires that the Hessian matrix

$$\mathbf{D}^2 f(\mathbf{x}^*) = \left( \frac{\partial^2 f(\mathbf{x}^*)}{\partial x_i \partial x_j} \right)$$

must satisfy the condition that  $\mathbf{h} \mathbf{D}^2 f(\mathbf{x}^*) \mathbf{h}' \leq 0$  for all vectors  $\mathbf{h}$ . (The prime indicates the transpose operation.) Geometrically, the requirement that the Hessian matrix is

negative semi-definite means that the production function must be locally concave in the neighborhood of an optimal choice. Formerly, we have the following proposition.

**Proposition 3.3.2** *Suppose that  $f(x)$  is differentiable and concave on  $\mathbb{R}_+^L$  and  $(p, w) > 0$ . If  $x > 0$  satisfies the first-order conditions given in (3.4), then  $x$  is (globally) profit maximizing production plan at prices  $(p, w)$ .*

**Remark 3.3.2** The strict concavity of  $f(\mathbf{x})$  can be checked by verifying if the leading principal minors of the Hessian must alternate in sign, i.e.,

$$\begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} > 0,$$

$$\begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} < 0,$$

and so on, where  $f_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ . This algebraic condition is useful for checking second-order conditions.

**Example 3.3.1 (The Profit Function for Cobb-Douglas Technology)** Consider the problem of maximizing profits for the production function of the form  $f(x) = x^a$  where  $a > 0$ . The first-order condition is

$$pax^{a-1} = w,$$

and the second-order condition reduces to

$$pa(a-1)x^{a-2} \leq 0.$$

The second-order condition can only be satisfied when  $a \leq 1$ , which means that the production function must have constant or decreasing returns to scale for competitive profit maximization to be meaningful.

If  $a = 1$ , the first-order condition reduces to  $p = w$ . Hence, when  $w = p$  any value of  $x$  is a profit-maximizing choice. When  $a < 1$ , we use the first-order condition to solve for the factor demand function

$$x(p, w) = \left( \frac{w}{ap} \right)^{\frac{1}{a-1}}.$$

The supply function is given by

$$y(p, w) = f(x(p, w)) = \left(\frac{w}{ap}\right)^{\frac{a}{a-1}},$$

and the profit function is given by

$$\pi(p, w) = py(p, w) - wx(p, w) = w \left(\frac{1-a}{a}\right) \left(\frac{w}{ap}\right)^{\frac{1}{a-1}}.$$

### 3.3.5 Properties of Net Supply Functions

In this section, we show that the net supply functions are the solutions to the profit maximization problem that in fact have imposed certain restrictions on the behavior of the demand and supply functions.

**Proposition 3.3.3** *Net output functions  $y(p)$  are homogeneous of degree zero, i.e.,  $y(tp) = y(p)$  for all  $t > 0$ .*

Proof. It is easy to see that if we multiply all of the prices by some positive number  $t$ , the production plan that maximizes profits will not change. Hence, we must have  $y(tp) = y(p)$  for all  $t > 0$ .

**Proposition 3.3.4 (Negative Definiteness of Substitution Matrix)** *Let  $y = f(x)$  be a twice differentiable and strictly concave single output production function, and let  $x(p, w)$  be the input demand function. Then, the substitution matrix*

$$Dx(p, w) = \left[ \frac{\partial x_i(p, w)}{\partial w_j} \right]$$

*is symmetric negative definite.*

Proof. Without loss of generality, we normalize  $p = 1$ . Then the first-order conditions for profit maximization are

$$\mathbf{D}f(\mathbf{x}(\mathbf{w})) - \mathbf{w} \equiv \mathbf{0}.$$

If we differentiate with respect to  $\mathbf{w}$ , we get

$$\mathbf{D}^2 f(\mathbf{x}(\mathbf{w})) \mathbf{D}\mathbf{x}(\mathbf{w}) - \mathbf{I} \equiv \mathbf{0}.$$

Solving this equation for the substitution matrix, we find

$$\mathbf{D}\mathbf{x}(\mathbf{w}) \equiv [\mathbf{D}^2 f(\mathbf{x}(\mathbf{w}))]^{-1}$$

Recall that the second-order condition for (strict) profit maximization is that the Hessian matrix is a symmetric negative definite matrix. It is a standard result of linear algebra that the inverse of a symmetric negative definite matrix is a symmetric negative definite matrix. Then,  $\mathbf{D}^2 f(\mathbf{x}(\mathbf{w}))$  is a symmetric negative definite matrix, and thus the substitution matrix  $\mathbf{D}\mathbf{x}(\mathbf{w})$  is a symmetric negative definite matrix. This means that the substitution matrix itself must be a symmetric, negative definite matrix.

**Remark 3.3.3** Note that since  $Dx(p, w)$  symmetric, negative definite, we particularly have:

- (1)  $\partial x_i / \partial w_i < 0$ , for  $i = 1, 2, \dots, n$  since the diagonal entries of a negative definite matrix must be negative.
- (2)  $\partial x_i / \partial w_j = \partial x_j / \partial w_i$  by the symmetry of the matrix.

### 3.3.6 Weak Axiom of Profit Maximization

In this subsection we examine the consequences of profit-maximizing behavior. Suppose that we have are given a list of observed price vectors  $\mathbf{p}^t$ , and the associated net output vectors  $\mathbf{y}^t$ , for  $t = 1, \dots, T$ . We refer to this collection as the *data*. In terms of the net supply functions we described before, the data are just  $(\mathbf{p}^t, \mathbf{y}(\mathbf{p}^t))$  for some observations  $t = 1, \dots, T$ . If the firm is maximizing profits, then the observed net output choice at price  $\mathbf{p}^t$  must have a level of profit at least as great as the profit at any other net output the firm could have chosen. Thus, a *necessary* condition for profit maximization is

$$\mathbf{p}^t \mathbf{y}^t \geq \mathbf{p}^t \mathbf{y}^s \quad \text{for all } t \quad \text{and } s = 1, \dots, T.$$

We will refer to this condition as the **Weak Axiom of Profit Maximization** (WTAPM).

In Figure 3.4A we have drawn two observations that *violate* WAPM, while Figure 3.4B depicts two observations that *satisfy* WAPM.

WAPM is a simple, but very useful, condition; let us derive some of its consequences. Fix two observations  $t$  and  $s$ , and write WAPM for each one. We have

$$\begin{aligned} \mathbf{p}^t (\mathbf{y}^t - \mathbf{y}^s) &\geq 0 \\ -\mathbf{p}^s (\mathbf{y}^t - \mathbf{y}^s) &\geq 0. \end{aligned}$$

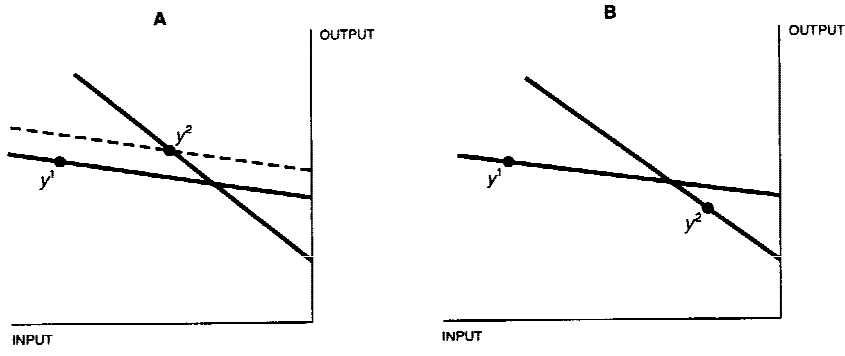


Figure 3.4: WAPM. Panel A shows two observations that violate WAPM since  $p^1 y^2 > p^1 y^1$ . Panel B shows two observations that satisfy WAPM.

Adding these two inequalities gives us

$$(\mathbf{p}^t - \mathbf{p}^s)(\mathbf{y}^t - \mathbf{y}^s) \geq 0.$$

Letting  $\Delta \mathbf{p} = (\mathbf{p}^t - \mathbf{p}^s)$  and  $\Delta \mathbf{y} = (\mathbf{y}^t - \mathbf{y}^s)$ , we can rewrite this expression as

$$\Delta \mathbf{p} \Delta \mathbf{y} \geq 0.$$

In other words, *the inner product of a vector of price changes with the associated vector of changes in net outputs must be nonnegative.*

### 3.3.7 Recoverability

Does WAPM exhaust all of the implications of profit-maximizing behavior, or are there other useful conditions implied by profit maximization? One way to answer this question is to try to construct a technology that generates the observed behavior  $(\mathbf{p}^t, \mathbf{y}^t)$  as profit-maximizing behavior. If we can find such a technology for *any* set of data that satisfy WAPM, then WAPM must indeed exhaust the implications of profit-maximizing behavior. We refer to the operation of constructing a technology consistent with the observed choices as the operation of **recoverability**.

We will show that if a set of data satisfies WAPM it is always possible to find a technology for which the observed choices are profit-maximizing choices. In fact, it is always possible to find a production set  $Y$  that is closed and convex. The remainder of this subsection will sketch the proof of this assertion. Formerly, we have

**Proposition 3.3.5** *For any set of data satisfies WAPM, there is a convex and closed production set such that the observed choices are profit-maximizing choices.*

Proof. We want to construct a convex and closed production set that will generate the observed choices  $(\mathbf{p}^t, \mathbf{y}^t)$  as profit-maximizing choices. We can actually construct two such production sets, one that serves as an “*inner bound*” to the true technology and one that serves as an “*outer bound*.” We start with the inner bound.

Suppose that the true production set  $Y$  is convex and monotonic. Since  $Y$  must contain  $\mathbf{y}^t$  for  $t = 1, \dots, T$ , it is natural to take the inner bound to be the smallest convex, monotonic set that contains  $\mathbf{y}^1, \dots, \mathbf{y}^T$ . This set is called the convex, monotonic **hull** of the points  $\mathbf{y}^1, \dots, \mathbf{y}^T$  and is denoted by

$$YI = \text{convex, monotonic hull of } \{\mathbf{y}^t : t = 1, \dots, T\}$$

The set  $YI$  is depicted in Figure 3.5.

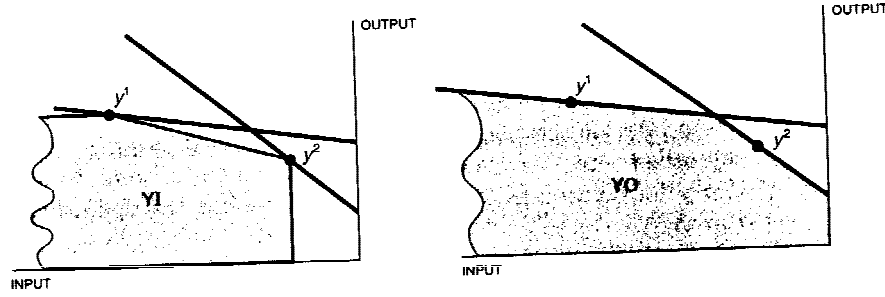


Figure 3.5: The set of  $YI$  and  $YO$ .

It is easy to show that for the technology  $YI$ ,  $\mathbf{y}^t$  is a profit-maximizing choice prices  $\mathbf{p}^t$ . All we have to do is to check that for all  $t$ ,

$$\mathbf{p}^t \mathbf{y}^t \geq \mathbf{p}^t \mathbf{y} \text{ for all } \mathbf{y} \text{ in } YI.$$

Suppose that this is not the case. Then for some observation  $t$ ,  $\mathbf{p}^t \mathbf{y}^t < \mathbf{p}^t \mathbf{y}$  for some  $\mathbf{y}$  in  $YI$ . But inspecting the diagram shows that there must then exist some observation  $s$  such that  $\mathbf{p}^t \mathbf{y}^t < \mathbf{p}^t \mathbf{y}^s$ . But this inequality violates WAPM.

Thus the set  $YI$  **rationalizes** the observed behavior in the sense that it is one possible technology that could have generated that behavior. It is not hard to see that  $YI$  must be

contained in any convex technology that generated the observed behavior: if  $Y$  generated the observed behavior and it is convex, then it must contain the observed choices  $\mathbf{y}^t$  and the convex hull of these points is the smallest such set. In this sense,  $YI$  gives us an “inner bound” to the true technology that generated the observed choices.

It is natural to ask if we can find an outer bound to this “true” technology. That is, can we find a set  $YO$  that is guaranteed to contain any technology that is consistent with the observed behavior?

The trick to answering this question is to rule out all of the points that couldn’t possibly be in the true technology and then take everything that is left over. More precisely, let us define  $NOTY$  by

$$NOTY = \{\mathbf{y}: \mathbf{p}^t \mathbf{y} > \mathbf{p}^t \mathbf{y}^t \text{ for some } t\}.$$

$NOTY$  consists of all those net output bundles that yield higher profits than some observed choice. If the firm is a profit maximizer, such bundles couldn’t be technologically feasible; otherwise they would have been chosen. Now as our outer bound to  $Y$  we just take the complement of this set:

$$YO = \{\mathbf{y}: \mathbf{p}^t \mathbf{y} \leq \mathbf{p}^t \mathbf{y}^t \text{ for all } t = 1, \dots, T\}.$$

The set  $YO$  is depicted in Figure 3.5B.

In order to show that  $YO$  rationalizes the observed behavior we must show that the profits at the observed choices are at least as great as the profits at any other  $\mathbf{y}$  in  $YO$ . Suppose not. Then there is some  $\mathbf{y}^t$  such that  $\mathbf{p}^t \mathbf{y}^t < \mathbf{p}^t \mathbf{y}$  for some  $\mathbf{y}$  in  $YO$ . But this contradicts the definition of  $YO$  given above. It is clear from the construction of  $YO$  that it must contain any production set consistent with the data  $(\mathbf{y}^t)$ . Hence,  $YO$  and  $YI$  form the tightest inner and outer bounds to the true production set that generated the data.

### 3.4 Profit Function

Given any production set  $Y$ , we have seen how to calculate the profit function,  $\pi(\mathbf{p})$ , which gives us the maximum profit attainable at prices  $\mathbf{p}$ . The profit function possesses

several important properties that follow directly from its definition. These properties are very useful for analyzing profit-maximizing behavior.

### 3.4.1 Properties of the Profit Function

The properties given below follow solely from the assumption of profit maximization. No assumptions about convexity, monotonicity, or other sorts of regularity are necessary.

**Proposition 3.4.1 (Properties of the Profit Function)** *The following has the following properties:*

- (1) *Nondecreasing in output prices, nonincreasing in input prices. If  $p'_i \geq p_i$  for all outputs and  $p'_j \leq p_j$ , for all inputs, then  $\pi(\mathbf{p}') \geq \pi(\mathbf{p})$ .*
- (2) *Homogeneous of degree 1 in  $\mathbf{p}$ .  $\pi(t\mathbf{p}) = t\pi(\mathbf{p})$  for all  $t \geq 0$ .*
- (3) *Convex in  $\mathbf{p}$ . Let  $\mathbf{p}'' = t\mathbf{p} + (1-t)\mathbf{p}'$  for  $0 \leq t \leq 1$ . Then  $\pi(\mathbf{p}'') \leq t\pi(\mathbf{p}) + (1-t)\pi(\mathbf{p}')$ .*
- (4) *Continuous in  $\mathbf{p}$ . The function  $\pi(\mathbf{p})$  is continuous, at least when  $\pi(\mathbf{p})$  is well-defined and  $p_i > 0$  for  $i = 1, \dots, n$ .*

PROOF.

1. Let  $\mathbf{y}$  be a profit-maximizing net output vector at  $\mathbf{p}$ , so that  $\pi(\mathbf{p}) = \mathbf{p}\mathbf{y}$  and let  $\mathbf{y}'$  be a profit-maximizing net output vector at  $\mathbf{p}'$  so that  $\pi(\mathbf{p}') = \mathbf{p}'\mathbf{y}'$ . Then by definition of profit maximization we have  $\mathbf{p}'\mathbf{y}' \geq \mathbf{p}'\mathbf{y}$ . Since  $p'_i \geq p_i$  for all  $i$  for which  $y_i \geq 0$  and  $p'_i \leq p_i$  for all  $i$  for which  $y_i \leq 0$ , we also have  $\mathbf{p}'\mathbf{y} \geq \mathbf{p}\mathbf{y}$ . Putting these two inequalities together, we have  $\pi(\mathbf{p}') = \mathbf{p}'\mathbf{y}' \geq \mathbf{p}\mathbf{y} = \pi(\mathbf{p})$ , as required.
2. Let  $\mathbf{y}$  be a profit-maximizing net output vector at  $\mathbf{p}$ , so that  $\mathbf{p}\mathbf{y} \geq \mathbf{p}\mathbf{y}'$  for all  $\mathbf{y}'$  in  $Y$ . It follows that for  $t \geq 0$ ,  $t\mathbf{p}\mathbf{y} \geq t\mathbf{p}\mathbf{y}'$  for all  $\mathbf{y}'$  in  $Y$ . Hence  $\mathbf{y}$  also maximizes profits at prices  $t\mathbf{p}$ . Thus  $\pi(t\mathbf{p}) = t\mathbf{p}\mathbf{y} = t\pi(\mathbf{p})$ .
3. Let  $\mathbf{y}$  maximize profits at  $\mathbf{p}$ ,  $\mathbf{y}'$  maximize profits at  $\mathbf{p}'$ , and  $\mathbf{y}''$  maximize profits at  $\mathbf{p}''$ . Then we have

$$\pi(\mathbf{p}'') = \mathbf{p}''\mathbf{y}'' = (t\mathbf{p} + (1-t)\mathbf{p}')\mathbf{y}'' = t\mathbf{p}\mathbf{y}'' + (1-t)\mathbf{p}'\mathbf{y}''. \quad (3.1)$$



By the definition of profit maximization, we know that

$$\begin{aligned} t\mathbf{p}\mathbf{y}'' &\leq t\mathbf{p}\mathbf{y} = t\pi(\mathbf{p}) \\ (1-t)\mathbf{p}'\mathbf{y}'' &\leq (1-t)\mathbf{p}'\mathbf{y}' = (1-t)\pi(\mathbf{p}'). \end{aligned}$$

Adding these two inequalities and using (3.1), we have

$$\pi(\mathbf{p}'') \leq t\pi(\mathbf{p}) + (1-t)\pi(\mathbf{p}'),$$

as required.

4. The continuity of  $\pi(\mathbf{p})$  follows from the Theorem of the Maximum.

### 3.4.2 Deriving Net Supply Functions from Profit Function

If we are given the net supply function  $\mathbf{y}(\mathbf{p})$ , it is easy to calculate the profit function. We just substitute into the definition of profits to find  $\pi(\mathbf{p}) = \mathbf{p}\mathbf{y}(\mathbf{p})$ . Suppose that instead we are given the profit function and are asked to find the net supply functions. How can that be done? It turns out that there is a very simple way to solve this problem: just differentiate the profit function. The proof that this works is the content of the next proposition.

**Proposition 3.4.2 (Hotelling's lemma.)** *Let  $y_i(\mathbf{p})$  be the firm's net supply function for good  $i$ . Then*

$$y_i(\mathbf{p}) = \frac{\partial \pi(\mathbf{p})}{\partial p_i} \quad \text{for } i = 1, \dots, n,$$

*assuming that the derivative exists and that  $p_i > 0$ .*

PROOF. Suppose  $(\mathbf{y}^*)$  is a profit-maximizing net output vector at prices  $(\mathbf{p}^*)$ . Then define the function

$$g(\mathbf{p}) = \pi(\mathbf{p}) - \mathbf{p}\mathbf{y}^*.$$

Clearly, the profit-maximizing production plan at prices  $\mathbf{p}$  will always be at least as profitable as the production plan  $\mathbf{y}^*$ . However, the plan  $\mathbf{y}^*$  will be a profit-maximizing plan at prices  $\mathbf{p}^*$ , so the function  $g$  reaches a minimum value of 0 at  $\mathbf{p}^*$ . The assumptions on prices imply this is an interior minimum.

The first-order conditions for a minimum then imply that

$$\frac{\partial g(\mathbf{p}^*)}{\partial p_i} = \frac{\partial \pi(\mathbf{p}^*)}{\partial p_i} - y_i^* = 0 \quad \text{for } i = 1, \dots, n.$$

Since this is true for all choices of  $\mathbf{p}^*$ , the proof is done.

**Remark 3.4.1** Again, we can prove this derivative property of the profit function by applying Envelope Theorem:

$$\left. \frac{d\pi(p)}{dp_i} = \frac{\partial py(p)}{\partial p_i} \right|_{x=x(a)} = y_i.$$

This expression says that the derivative of  $\pi$  with respect to  $a$  is given by the partial derivative of  $f$  with respect to  $p_i$ , holding  $x$  fixed at the optimal choice. This is the meaning of the vertical bar to the right of the derivative.

## 3.5 Cost Minimization

An important implication of the firm choosing a profit-maximizing production plan is that there is no way to produce the same amounts of outputs at a lower total input cost. Thus, cost minimization is a necessary condition for profit maximization. This observation motivates us to an independent study of the firm's cost minimization. The problem is of interest for several reasons. First, it leads us to a number of results and constructions that are technically very useful. Second, as long as the firm is a price taker in its input market, the results flowing from the cost minimization continue to be valid whether or not the output market is competitive and so whether or not the firm takes the output price as given as. Third, when the production set exhibits nondecreasing returns to scale, the cost function and optimizing vectors of the cost minimization problem, which keep the levels of outputs fixed, are better behaved than the profit function.

To be concrete, we focus our analysis on the single-output case. We assume throughout that firms are perfectly competitive on their input markets and therefore they face fixed prices. Let  $w = (w_1, w_2, \dots, w_n) \geq 0$  be a vector of prevailing market prices at which the firm can buy inputs  $x = (x_1, x_2, \dots, x_n)$ .

### 3.5.1 First-Order Conditions of Cost Minimization

Let us consider the problem of finding a cost-minimizing way to produce a given level of output:

$$\begin{aligned} \min_x \quad & \mathbf{w}\mathbf{x} \\ \text{such that} \quad & f(\mathbf{x}) \geq y \end{aligned}$$

We analyze this constrained minimization problem using the Lagrangian function:

$$\mathcal{L}(\lambda, \mathbf{x}) = \mathbf{w}\mathbf{x} - \lambda(f(\mathbf{x}) - y)$$

where production function  $f$  is assumed to be differentiable and  $\lambda$  is the Lagrange multiplier. The first-order conditions characterizing an interior solution  $\mathbf{x}^*$  are

$$w_i - \lambda \frac{\partial f(\mathbf{x}^*)}{\partial x_i} = 0, \quad i = 1, \dots, n \quad (3.5)$$

$$f(\mathbf{x}^*) = y \quad (3.6)$$

or in vector notation, the condition can be written as

$$\mathbf{w} = \lambda \mathbf{D}f(\mathbf{x}^*).$$

We can interpret these first-order conditions by dividing the  $j^{th}$  condition by the  $i^{th}$  condition to get

$$\frac{w_i}{w_j} = \frac{\frac{\partial f(\mathbf{x}^*)}{\partial x_i}}{\frac{\partial f(\mathbf{x}^*)}{\partial x_j}} \quad i, j = 1, \dots, n, \quad (3.7)$$

which means the marginal rate of technical substitution of factor  $i$  for factor  $j$  equals the economic rate of substitution factor  $i$  for factor  $j$  at the cost minimizing input bundle.

This first-order condition can also be represented graphically. In Figure 3.6, the curved lines represent iso-quants and the straight lines represent constant cost curves. When  $y$  is fixed, the problem of the firm is to find a cost-minimizing point on a given iso-quant. It is clear that such a point will be characterized by the tangency condition that the slope of the constant cost curve must be equal to the slope of the iso-quant.

Again, the conditions are valid only for interior operating positions: they must be modified if a cost minimization point occurs on the boundary. The appropriate conditions turn out to be

$$\lambda \frac{\partial f(\mathbf{x}^*)}{\partial x_i} - w_i \leq 0 \quad \text{with equality if } x_i > 0, \quad i = 1, 2, \dots, n \quad (3.8)$$

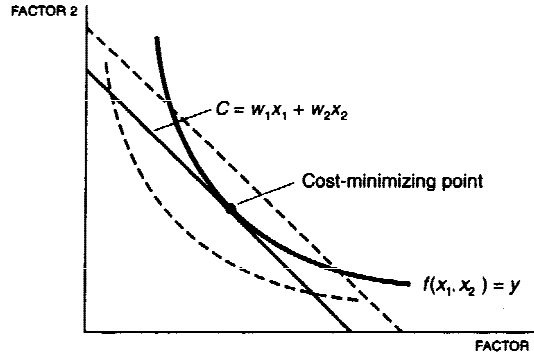


Figure 3.6: Cost minimization. At a point that minimizes costs, the isoquant must be tangent to the constant cost line.

**Remark 3.5.1** It is known that a continuous function achieves a minimum and a maximum value on a closed and bounded set. The objective function  $\mathbf{w}\mathbf{x}$  is certainly a continuous function and the set  $V(y)$  is a closed set by hypothesis. All that we need to establish is that we can restrict our attention to a bounded subset of  $V(y)$ . But this is easy. Just pick an arbitrary value of  $\mathbf{x}$ , say  $\mathbf{x}'$ . Clearly the minimal cost factor bundle must have a cost less than  $\mathbf{w}\mathbf{x}'$ . Hence, we can restrict our attention to the subset  $\{\mathbf{x} \text{ in } V(y): \mathbf{w}\mathbf{x} \leq \mathbf{w}\mathbf{x}'\}$ , which will certainly be a bounded subset, as long as  $w > 0$ . Thus the cost minimizing input bundle always exists.

### 3.5.2 Sufficiency of First-Order Conditions for Cost Minimization

Again, like consumer's constrained optimization problem, the above first-order conditions are merely necessary conditions for a local optimum. However, these necessary first-order conditions are in fact sufficient for a global optimum when a production function is quasi-concave, which is formerly stated in the following proposition.

**Proposition 3.5.1** *Suppose that  $f(x) : \mathbb{R}_+^n \rightarrow \mathbb{R}$  is differentiable and quasi-concave on  $\mathbb{R}_+^n$  and  $w > 0$ . If  $(x, \lambda) > 0$  satisfies the first-order conditions given in (3.5) and (3.6), then  $x$  solves the firm's cost minimization problem at prices  $w$ .*

**PROOF.** Since  $f(x) : \mathbb{R}_+^n \rightarrow \mathbb{R}$  is differentiable and quasi-concave, the input requirement set  $V(y) = \{x : f(x) \geq y\}$  is a convex and closed set. Further the object function  $w\mathbf{x}$  is

convex and continuous, then by the Kuhn-Tucker theorem, the first-order conditions are sufficient for the constrained minimization problem.

Similarly, the strict quasi-concavity of  $f$  can be checked by verifying if the naturally ordered principal minors of the bordered Hessian alternative in sign, i.e.,

$$\begin{vmatrix} 0 & f_1 & f_2 \\ f_1 & f_{11} & f_{12} \\ f_2 & f_{21} & f_{22} \end{vmatrix} > 0,$$

$$\begin{vmatrix} 0 & f_1 & f_2 & f_3 \\ f_1 & f_{11} & f_{12} & f_{13} \\ f_2 & f_{21} & f_{22} & f_{23} \\ f_3 & f_{31} & f_{32} & f_{33} \end{vmatrix} < 0,$$

and so on, where  $f_i = \frac{\partial f}{\partial x_i}$  and  $f_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ .

For each choice of  $\mathbf{w}$  and  $y$  there will be some choice of  $\mathbf{x}^*$  that minimizes the cost of producing  $y$  units of output. We will call the function that gives us this optimal choice the **conditional input demand function** and write it as  $\mathbf{x}(\mathbf{w}, y)$ . Note that conditional factor demands depend on the level of output produced as well as on the factor prices. The **cost function** is the minimal cost at the factor prices  $\mathbf{w}$  and output level  $y$ ; that is.  $c(\mathbf{w}, y) = \mathbf{w}\mathbf{x}(\mathbf{w}, y)$ .

**Example 3.5.1 (Cost Function for the Cobb-Douglas Technology)** Consider the cost minimization problem

$$\begin{aligned} c(w, y) &= \min_{x_1, x_2} w_1 x_1 + w_2 x_2 \\ &\text{such that } Ax_1^a x_2^b = y. \end{aligned}$$

Solving the constraint for  $x_2$ , we see that this problem is equivalent to

$$\min_{x_1} w_1 x_1 + w_2 A^{-\frac{1}{b}} y^{\frac{1}{b}} x_1^{-\frac{a}{b}}.$$

The first-order condition is

$$w_1 - \frac{a}{b} w_2 A^{-\frac{1}{b}} y^{\frac{1}{b}} x_1^{-\frac{a+b}{b}} = 0,$$

which gives us the conditional input demand function for factor 1:

$$x_1(w_1, w_2, y) = A^{-\frac{1}{a+b}} \left[ \frac{aw_1}{bw_2} \right]^{\frac{b}{a+b}} y^{\frac{1}{a+b}}.$$

The other conditional input demand function is

$$x_2(w_1, w_2, y) = A^{-\frac{1}{a+b}} \left[ \frac{aw_2}{bw_1} \right]^{\frac{a}{a+b}} y^{\frac{1}{a+b}}.$$

The cost function is thus

$$\begin{aligned} c(w_1, w_2, y) &= w_1 x_1(w_1, w_2, y) + w_2 x_2(w_1, w_2, y) \\ &= A^{\frac{-1}{a+b}} \left[ \left( \frac{a}{b} \right)^{\frac{b}{a+b}} \left( \frac{a}{b} \right)^{\frac{-a}{a+b}} \right] w_1^{\frac{a}{a+b}} w_2^{\frac{b}{a+b}} y^{\frac{1}{a+b}}. \end{aligned}$$

When  $A = 1$  and  $a + b = 1$  (constant returns to scale), we particularly have

$$c(w_1, w_2, y) = K w_1^a w_2^{1-a} y,$$

where  $K = a^{-a}(1-a)^{a-1}$ .

**Example 3.5.2 (The Cost function for the CES technology)** Suppose that  $f(x_1, x_2) = (x_1^\rho + x_2^\rho)^{\frac{1}{\rho}}$ . What is the associated cost function? The cost minimization problem is

$$\begin{aligned} \min w_1 x_1 + w_2 x_2 \\ \text{such that } x_1^\rho + x_2^\rho = y^\rho \end{aligned}$$

The first-order conditions are

$$\begin{aligned} w_1 - \lambda \rho x_1^{\rho-1} &= 0 \\ w_2 - \lambda \rho x_2^{\rho-1} &= 0 \\ x_1^\rho + x_2^\rho &= y. \end{aligned}$$

Solving the first two equations for  $x_1^\rho$  and  $x_2^\rho$ , we have

$$x_1^\rho = w_1^{\frac{\rho}{\rho-1}} (\lambda \rho)^{\frac{-\rho}{\rho-1}} \quad (3.9)$$

$$x_2^\rho = w_2^{\frac{\rho}{\rho-1}} (\lambda \rho)^{\frac{-\rho}{\rho-1}}. \quad (3.10)$$

Substitute this into the production function to find

$$(\lambda \rho)^{\frac{-\rho}{\rho-1}} \left( w_1^{\frac{\rho}{\rho-1}} + w_2^{\frac{\rho}{\rho-1}} \right) = y^\rho.$$

Solve this for  $(\lambda \rho)^{\frac{-\rho}{\rho-1}}$  and substitute into equations (3.9) and (3.10). This gives us the conditional input demand functions

$$\begin{aligned} x_1(w_1, w_2, y) &= w_1^{\frac{1}{\rho-1}} \left[ w_1^{\frac{\rho}{\rho-1}} + w_2^{\frac{\rho}{\rho-1}} \right]^{-\frac{1}{\rho}} y \\ x_2(w_1, w_2, y) &= w_2^{\frac{1}{\rho-1}} \left[ w_1^{\frac{\rho}{\rho-1}} + w_2^{\frac{\rho}{\rho-1}} \right]^{-\frac{1}{\rho}} y. \end{aligned}$$

Substituting these functions into the definition of the cost function yields

$$\begin{aligned}
c(w_1, w_2, y) &= w_1 x_1(w_1, w_2, y) + w_2 x_2(w_1, w_2, y) \\
&= y \left[ w_1^{\frac{\rho}{\rho-1}} + w_2^{\frac{\rho}{\rho-1}} \right] \left[ w_1^{\frac{\rho}{\rho-1}} + w_2^{\frac{\rho}{\rho-1}} \right]^{-\frac{1}{\rho}} \\
&= y \left[ w_1^{\frac{\rho}{\rho-1}} + w_2^{\frac{\rho}{\rho-1}} \right]^{\frac{\rho-1}{\rho}}.
\end{aligned}$$

This expression looks a bit nicer if we set  $r = \rho/(\rho - 1)$  and write

$$c(w_1, w_2, y) = y[w_1^r + w_2^r]^{\frac{1}{r}}.$$

Note that this cost function has the same form as the original CES production function with  $r$  replacing  $\rho$ . In the general case where

$$f(x_1, x_2) = [(a_1 x_1)^\rho + (a_2 x_2)^\rho]^{\frac{1}{\rho}},$$

similar computations can be done to show that

$$c(w_1, w_2, y) = [(w_1/a_1)^r + (w_2/a_2)^r]^{\frac{1}{r}} y.$$

**Example 3.5.3 (The Cost function for the Leontief technology)** Suppose  $f(x_1, x_2) = \min\{ax_1, bx_2\}$ . Since we know that the firm will not waste any input with a positive price, the firm must operate at a point where  $y = ax_1 = bx_2$ . Hence, if the firm wants to produce  $y$  units of output, it must use  $y/a$  units of good 1 and  $y/b$  units of good 2 no matter what the input prices are. Hence, the cost function is given by

$$c(w_1, w_2, y) = \frac{w_1 y}{a} + \frac{w_2 y}{b} = y \left( \frac{w_1}{a} + \frac{w_2}{b} \right).$$

**Example 3.5.4 (The cost function for the linear technology)** Suppose that  $f(x_1, x_2) = ax_1 + bx_2$ , so that factors 1 and 2 are perfect substitutes. What will the cost function look like? Since the two goods are perfect substitutes, the firm will use whichever is cheaper. Hence, the cost function will have the form  $c(w_1, w_2, y) = \min\{w_1/a, w_2/b\}y$ .

In this case the answer to the cost-minimization problem typically involves a boundary solution: one of the two factors will be used in a zero amount. It is easy to see the answer to this particular problem by comparing the relative steepness of the isocost line and isoquant curve. If  $\frac{a_1}{a_2} < \frac{w_1}{w_2}$ , the firm only uses  $x_2$  and the cost function is given by  $c(w_1, w_2, y) = w_2 x_2 = w_2 \frac{y}{a_2}$ . If  $\frac{a_1}{a_2} > \frac{w_1}{w_2}$ , the firm only uses  $x_1$  and the cost function is given by  $c(w_1, w_2, y) = w_1 x_1 = w_1 \frac{y}{a_1}$ .

## 3.6 Cost Functions

The cost function measures the minimum cost of producing a given level of output for some fixed factor prices. As such it summarizes information about the technological choices available to the firms. It turns out that the behavior of the cost function can tell us a lot about the nature of the firm's technology. In the following we will first investigate the behavior of the cost function  $c(\mathbf{w}, y)$  with respect to its price and quantity arguments. We then define a few related functions, namely the average and the marginal cost functions.

### 3.6.1 Properties of Cost Functions

You may have noticed some similarities here with consumer theory. These similarities are in fact exact when one compares the cost function with the expenditure function. Indeed, consider their definitions.

(1) Expenditure Function:  $e(p, u) \equiv \min_{x \in \mathbb{R}_+^n} px$  such that  $u(x) \geq u$

(2) Cost Function:  $c(w, y) \equiv \min_{x \in \mathbb{R}_+^n} wx$  such that  $f(x) \geq y$

Mathematically, the two optimization problems are identical. Consequently, for every theorem we proved about expenditure functions, there is an equivalent theorem for cost functions. We shall state these results here, but we do not need to prove them. Their proofs are identical to those given for the expenditure function.

**Proposition 3.6.1** [*Properties of the Cost Function.*] Suppose the production function  $f$  is continuous and strictly increasing. Then the cost function has the following properties:

- (1)  $c(\mathbf{w}, y)$  is nondecreasing in  $\mathbf{w}$ .
- (2)  $c(\mathbf{w}, y)$  is homogeneous of degree 1 in  $\mathbf{w}$ .
- (3)  $c(\mathbf{w}, y)$  is concave in  $\mathbf{w}$ .
- (4)  $c(\mathbf{w}, y)$  is continuous in  $\mathbf{w}$ , for  $w > 0$ .
- (5) For all  $w > 0$ ,  $c(w, \mathbf{y})$  is strictly increasing  $\mathbf{y}$ .



(6) *Shephard's lemma: If  $\mathbf{x}(\mathbf{w}, y)$  is the cost-minimizing bundle necessary to produce production level  $y$  at prices  $\mathbf{w}$ , then  $x_i(\mathbf{w}, y) = \frac{\partial c(\mathbf{w}, y)}{\partial w_i}$  for  $i = 1, \dots, n$  assuming the derivative exists and that  $x_i > 0$ .*

### 3.6.2 Properties of Conditional Input Demand

As solution to the firm's cost-minimization problem, the conditional input demand functions possess certain general properties. These are analogous to the properties of Hicksian compensation demands, so once again it is not necessary to repeat the proof.

**Proposition 3.6.2 (Negative Semi-Definite Substitution Matrix)** *The matrix of substitution terms  $(\partial x_j(\mathbf{w}, y)/\partial w_i)$  is negative semi-definite.*

Again since the substitution matrix is negative semi-definite, thus it is symmetric and has non-positive diagonal terms. We then particularly have

**Proposition 3.6.3 (Symmetric Substitution Terms)** *The matrix of substitution terms is symmetric, i.e.,*

$$\frac{\partial x_j(w, y)}{\partial w_i} = \frac{\partial^2 c(\mathbf{w}, y)}{\partial w_j \partial w_i} = \frac{\partial^2 c(\mathbf{w}, y)}{\partial w_i \partial w_j} = \frac{\partial x_i(\mathbf{w}, y)}{\partial w_j}.$$

**Proposition 3.6.4 (Negative Own-Substitution Terms)** *The compensated own-price effect is non-positive; that is, the input demand curves slope downward:*

$$\frac{\partial x_i(\mathbf{w}, y)}{\partial w_i} = \frac{\partial^2 c(\mathbf{w}, y)}{\partial w_i^2} \leq 0,$$

**Remark 3.6.1** Using the cost function, we can restate the firm's profit maximization problem as

$$\max_{y \geq 0} py - c(w, y). \quad (3.11)$$

The necessary first-order condition for  $y^*$  to be profit-maximizing is then

$$p - \frac{\partial c(w, y^*)}{\partial y} \leq 0, \quad \text{with equality if } y^* > 0. \quad (3.12)$$

In other words, at an interior optimum (i.e.,  $y^* > 0$ ), price equals marginal cost. If  $c(w, y)$  is convex in  $y$ , then the first-order condition (3.12) is also sufficient for  $y^*$  to be the firm's optimal output level.

### 3.6.3 Average and Marginal Costs

Let us consider the structure of the cost function. Note that the cost function can always be expressed simply as the value of the conditional factor demands.

$$c(\mathbf{w}, y) \equiv \mathbf{w}\mathbf{x}(\mathbf{w}, y)$$

In the short run, some of the factors of production are fixed at predetermined levels. Let  $\mathbf{x}_f$  be the vector of fixed factors,  $\mathbf{x}_v$ , the vector of variable factors, and break up  $\mathbf{w}$  into  $\mathbf{w} = (\mathbf{w}_v, \mathbf{w}_f)$ , the vectors of prices of the variable and the fixed factors. The short-run conditional factor demand functions will generally depend on  $\mathbf{x}_f$ , so we write them as  $\mathbf{x}_v(\mathbf{w}, y, \mathbf{x}_f)$ . Then the short-run cost function can be written as

$$c(\mathbf{w}, y, \mathbf{x}_f) = \mathbf{w}_v \mathbf{x}_v(\mathbf{w}, y, \mathbf{x}_f) + \mathbf{w}_f \mathbf{x}_f.$$

The term  $\mathbf{w}_v \mathbf{x}_v(\mathbf{w}, y, \mathbf{x}_f)$  is called **short-run variable cost** (SVC), and the term  $\mathbf{w}_f \mathbf{x}_f$  is the **fixed cost** (FC). We can define various derived cost concepts from these basic units:

$$\begin{aligned} \text{short-run total cost} &= STC = \mathbf{w}_v \mathbf{x}_v(\mathbf{w}, y, \mathbf{x}_f) + \mathbf{w}_f \mathbf{x}_f \\ \text{short-run average cost} &= SAC = \frac{c(\mathbf{w}, y, \mathbf{x}_f)}{y} \\ \text{short-run average variable cost} &= SAVC = \frac{\mathbf{w}_v \mathbf{x}_v(\mathbf{w}, y, \mathbf{x}_f)}{y} \\ \text{short-run average fixed cost} &= SAFC = \frac{\mathbf{w}_f \mathbf{x}_f}{y} \\ \text{short-run marginal cost} &= SMC = \frac{\partial c(\mathbf{w}, y, \mathbf{x}_f)}{\partial y}. \end{aligned}$$

When all factors are variable, the firm will optimize in the choice of  $\mathbf{x}_f$ . Hence, the long-run cost function only depends on the factor prices and the level of output as indicated earlier. We can express this long-run function in terms of the short-run cost function in the following way. Let  $\mathbf{x}_f(\mathbf{w}, y)$  be the optimal choice of the fixed factors, and let  $\mathbf{x}_v(\mathbf{w}, y) = \mathbf{x}_v(\mathbf{w}, y, \mathbf{x}_f(\mathbf{w}, y))$  be the long-run optimal choice of the variable factors. Then the long-run cost function can be written as

$$c(\mathbf{w}, y) = \mathbf{w}_v \mathbf{x}_v(\mathbf{w}, y) + \mathbf{w}_f \mathbf{x}_f(\mathbf{w}, y) = c(\mathbf{w}, y, \mathbf{x}_f(\mathbf{w}, y)).$$

Similarly, we can define the long-run average and marginal cost functions:

$$\begin{aligned}\text{long-run average cost} &= LAC = \frac{c(\mathbf{w}, y)}{y} \\ \text{long-run marginal cost} &= LMC = \frac{\partial c(\mathbf{w}, y)}{\partial y}.\end{aligned}$$

Notice that “long-run average cost” equals “long-run average variable cost” since all costs are variable in the long-run; “long-run fixed costs” are zero for the same reason.

**Example 3.6.1 (The short-run Cobb-Douglas cost functions)** Suppose the second factor in a Cobb-Douglas technology is restricted to operate at a level  $k$ . Then the cost-minimizing problem is

$$\begin{aligned}\min w_1 x_1 + w_2 k \\ \text{such that } y = x_1^a k^{1-a}.\end{aligned}$$

Solving the constraint for  $x_1$  as a function of  $y$  and  $k$  gives

$$x_1 = (y k^{a-1})^{\frac{1}{a}}.$$

Thus

$$c(w_1, w_2, y, k) = w_1 (y k^{a-1})^{\frac{1}{a}} + w_2 k.$$

The following variations can also be calculated:

$$\begin{aligned}\text{short-run average cost} &= w_1 \left(\frac{y}{k}\right)^{\frac{1-a}{a}} + \frac{w_2 k}{y} \\ \text{short-run average variable cost} &= w_1 \left(\frac{y}{k}\right)^{\frac{1-a}{a}} \\ \text{short-run average fixed cost} &= \frac{w_2 k}{y} \\ \text{short-run marginal cost} &= \frac{w_1}{a} \left(\frac{y}{k}\right)^{\frac{1-a}{a}}\end{aligned}$$

**Example 3.6.2 (Constant returns to scale and the cost function)** If the production function exhibits constant returns to scale, then it is intuitively clear that the cost function should exhibit costs that are linear in the level of output: if you want to produce twice as much output it will cost you twice as much. This intuition is verified in the following proposition:

**Proposition 3.6.5 (Constant returns to scale)** *If the production function exhibits constant returns to scale, the cost function may be written as  $c(\mathbf{w}, y) = yc(\mathbf{w}, 1)$ .*

PROOF. Let  $\mathbf{x}^*$  be a cheapest way to produce one unit of output at prices  $\mathbf{w}$  so that  $c(\mathbf{w}, 1) = \mathbf{w}\mathbf{x}^*$ . We want to show that  $c(\mathbf{w}, y) = y\mathbf{w}\mathbf{x}^* = yc(\mathbf{w}, 1)$ . Notice first that  $y\mathbf{x}^*$  is feasible to produce  $y$  since the technology is constant returns to scale. Suppose that it does not minimize cost; instead let  $\mathbf{x}'$  be the cost-minimizing bundle to produce  $y$  at prices  $\mathbf{w}$  so that  $\mathbf{w}\mathbf{x}' < \mathbf{w}y\mathbf{x}^*$ . Then  $\mathbf{w}\mathbf{x}'/y < \mathbf{w}\mathbf{x}^*$  and  $\mathbf{x}'/y$  can produce 1 since the technology is constant returns to scale. This contradicts the definition of  $\mathbf{x}^*$ .

Thus, if the technology exhibits constant returns to scale, then the average cost, the average variable cost, and the marginal cost functions are all the same.

### 3.6.4 The Geometry of Costs

Let us first examine the short-run cost curves. In this case, we will write the cost function simply as  $c(y)$ , which has two components: fixed costs and variable costs. We can therefore write short-run average cost as

$$SAC = \frac{c(\mathbf{w}, y, \mathbf{x}_f)}{y} = \frac{\mathbf{w}_f \mathbf{x}_f}{y} + \frac{\mathbf{w}_v \mathbf{x}_v(\mathbf{w}, y, \mathbf{x}_f)}{y} = SAFC + SAVC.$$

As we increase output, average variable costs may initially decrease if there is some initial region of economies of scale. However, it seems reasonable to suppose that the variable factors required will eventually increase by the law of diminishing marginal returns, as depicted in Figure 3.7. Average fixed costs must of course decrease with output, as indicated in Figure 3.7. Adding together the average variable cost curve and the average fixed costs gives us the  $U$ -shaped average cost curve in Figure 3.7. The initial decrease in average costs is due to the decrease in average fixed costs; the eventual increase in average costs is due to the increase in average variable costs. The level of output at which the average cost of production is minimized is sometimes known as the **minimal efficient scale**.

In the long run all costs are variable costs and the appropriate long-run average cost curve should also be  $U$ -shaped by the facts that variable costs usually exhibit increasing returns to scale at low level of production and ultimately exhibits decreasing returns to scale.

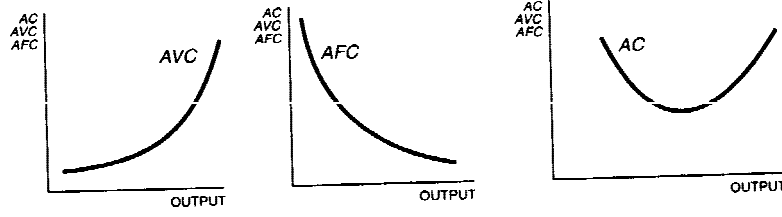


Figure 3.7: Average variable, average fixed, and average cost curves.

Let us now consider the marginal cost curve. What is its relationship to the average cost curve? Since

$$\frac{d}{dy} \left( \frac{c(y)}{y} \right) = \frac{yc'(y) - c(y)}{y^2} = \frac{1}{y} \left[ c'(y) - \frac{c(y)}{y} \right],$$

$$\frac{d}{dy} \left( \frac{c(y)}{y} \right) \leq 0 (\geq 0) \text{ if and only if } c'(y) - \frac{c(y)}{y} \leq 0 (\geq 0).$$

Thus, the average variable cost curve is decreasing when the marginal cost curve lies below the average variable cost curve, and it is increasing when the marginal cost curve lies above the average variable cost curve. It follows that average cost reach its minimum at  $y^*$  when the marginal cost curve passes through the average variable cost curve, i.e.,

$$c'(y^*) = \frac{c(y^*)}{y^*}.$$

**Remark 3.6.2** All of the analysis just discussed holds in both the long and the short run. However, if production exhibits constant returns to scale in the long run, so that the cost function is linear in the level of output, then average cost, average variable cost, and marginal cost are all equal to each other, which makes most of the relationships just described rather trivial.

### 3.6.5 Long-Run and Short-Run Cost Curves

Let us now consider the relationship between the long-run cost curves and the short-run cost curves. It is clear that the long-run cost curve must never lie above any short-run cost curve, since the short-run cost minimization problem is just a constrained version of the long-run cost minimization problem.

Let us write the long-run cost function as  $c(y) = c(y, z(y))$ . Here we have omitted the factor prices since they are assumed fixed, and we let  $z(y)$  be the cost-minimizing demand

for a single fixed factor. Let  $y^*$  be some given level of output, and let  $z^* = z(y^*)$  be the associated long-run demand for the fixed factor. The short-run cost,  $c(y, z^*)$ , must be at least as great as the long-run cost,  $c(y, z(y))$ , for all levels of output, and the short-run cost will equal the long-run cost at output  $y^*$ , so  $c(y^*, z^*) = c(y^*, z(y^*))$ . Hence, the long- and the short-run cost curves must be tangent at  $y^*$ .

This is just a geometric restatement of the envelope theorem. The slope of the long-run cost curve at  $y^*$  is

$$\frac{dc(y^*, z(y^*))}{dy} = \frac{\partial c(y^*, z^*)}{\partial y} + \frac{\partial c(y^*, z^*)}{\partial z} \frac{\partial z(y^*)}{\partial y}.$$

But since  $z^*$  is the *optimal* choice of the fixed factors at the output level  $y^*$ , we must have

$$\frac{\partial c(y^*, z^*)}{\partial z} = 0.$$

Thus, long-run marginal costs at  $y^*$  equal short-run marginal costs at  $(y^*, z^*)$ .

Finally, we note that if the long- and short-run cost curves are tangent, the long- and short-run *average* cost curves must also be tangent. A typical configuration is illustrated in Figure 3.8.

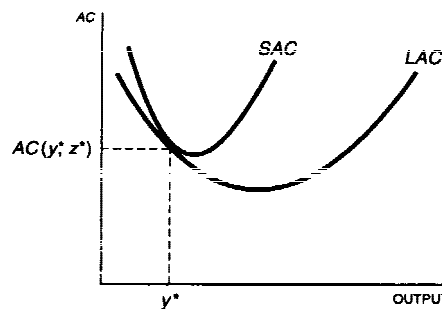


Figure 3.8: Long-run and short-run average cost curves. Note that the long-run and the short-run average cost curves must be tangent which implies that the long-run and short-run marginal cost must equal.

### 3.7 Duality in Production

In the last section we investigated the properties of the cost function. Given any technology, it is straightforward, at least in principle, to derive its cost function: we simply solve

the cost minimization problem.

In this section we show that this process can be reversed. Given a cost function we can “solve for” a technology that could have generated that cost function. This means that the cost function contains essentially the same information that the production function contains. Any concept defined in terms of the properties of the production function has a “dual” definition in terms of the properties of the cost function and vice versa. This general observation is known as the principle of **duality**.

### 3.7.1 Recovering a Production Set from a Cost Function

Given data  $(\mathbf{w}^t, \mathbf{x}^t, y^t)$ , define  $VO(y)$  as an “outer bound” to the true input requirement set  $V(y)$ :

$$VO(y) = \{\mathbf{x}: \mathbf{w}^t \mathbf{x} \geq \mathbf{w}^t \mathbf{x}^t \text{ for all } t \text{ such that } y^t \leq y\}.$$

It is straightforward to verify that  $VO(y)$  is a closed, monotonic, and convex technology. Furthermore, it contains any technology that could have generated the data  $(\mathbf{w}^t, \mathbf{x}^t, y^t)$  for  $t = 1, \dots, T$ .

If we observe choices for many different factor prices, it seems that  $VO(y)$  should “approach” the true input requirement set in some sense. To make this precise, let the factor prices vary over *all* possible price vectors  $\mathbf{w} \geq 0$ . Then the natural generalization of  $VO$  becomes

$$V^*(y) = \{x: \mathbf{w} \mathbf{x} \geq \mathbf{w} \mathbf{x}(\mathbf{w}, y) = c(\mathbf{w}, y) \text{ for all } \mathbf{w} \geq \mathbf{0}\}.$$

What is the relationship between  $V^*(y)$  and the true input requirement set  $V(y)$ ? Of course,  $V^*(y)$  clearly contain  $V(y)$ . In general,  $V^*(y)$  will strictly contain  $V(y)$ . For example, in Figure 3.9A we see that the shaded area cannot be ruled out of  $V^*(y)$  since the points in this area satisfy the condition that  $\mathbf{w} \mathbf{x} \geq c(\mathbf{w}, y)$ .

The same is true for Figure 3.9B. The cost function can only contain information about the *economically relevant* sections of  $V(y)$ , namely, those factor bundles that could actually be the solution to a cost minimization problem, i.e., that could actually be conditional factor demands.

However, suppose that our original technology is convex and monotonic. In this case  $V^*(y)$  will equal  $V(y)$ . This is because, in the convex, monotonic case, each point on

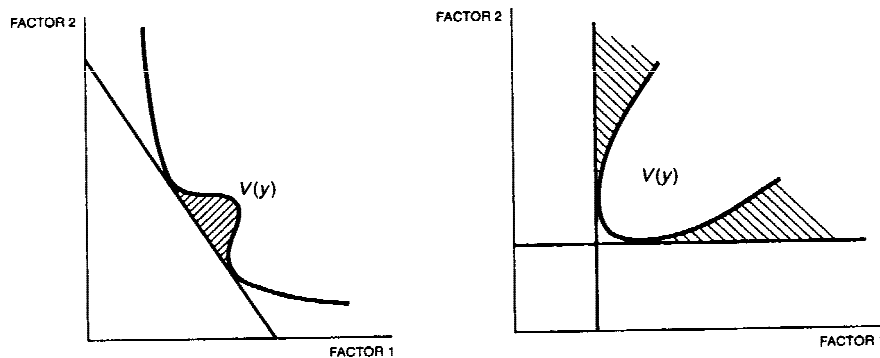


Figure 3.9: Relationship between  $V(y)$  and  $V^*(y)$ . In general,  $V^*(y)$  will strictly contain  $V(y)$ .

the boundary of  $V(y)$  is a cost-minimizing factor demand for *some* price vector  $\mathbf{w} \geq \mathbf{0}$ . Thus, the set of points where  $\mathbf{w}\mathbf{x} \geq c(\mathbf{w}, y)$  for all  $\mathbf{w} \geq \mathbf{0}$  will precisely describe the input requirement set. More formally:

**Proposition 3.7.1 (Equality of  $V(y)$  and  $V^*(y)$ )** *Suppose  $V(y)$  is a closed, convex, monotonic technology. Then  $V^*(y) = V(y)$ .*

PROOF (Sketch) We already know that  $V^*(y)$  contains  $V(y)$ , so we only have to show that if  $\mathbf{x}$  is in  $V^*(y)$  then  $\mathbf{x}$  must be in  $V(y)$ . Suppose that  $\mathbf{x}$  is not an element of  $V(y)$ . Then since  $V(y)$  is a closed convex set satisfying the monotonicity hypothesis, we can apply a version of the separating hyperplane theorem to find a vector  $\mathbf{w}^* \geq \mathbf{0}$  such that  $\mathbf{w}^*\mathbf{x} < \mathbf{w}^*\mathbf{z}$  for all  $\mathbf{z}$  in  $V(y)$ . Let  $\mathbf{z}^*$  be a point in  $V(y)$  that minimizes cost at the prices  $\mathbf{w}^*$ . Then in particular we have  $\mathbf{w}^*\mathbf{x} < \mathbf{w}^*\mathbf{z}^* = c(\mathbf{w}^*, y)$ . But then  $\mathbf{x}$  cannot be in  $V^*(y)$ , according to the definition of  $V^*(y)$ .

This proposition shows that if the original technology is convex and monotonic, then the cost function associated with the technology can be used to completely reconstruct the original technology. This is a reasonably satisfactory result in the case of convex and monotonic technologies, but what about less well-behaved cases? Suppose we start with some technology  $V(y)$ , possibly non-convex. We find its cost function  $c(\mathbf{w}, y)$  and then generate  $V^*(y)$ . We know from the above results that  $V^*(y)$  will not necessarily be equal to  $V(y)$ , unless  $V(y)$  happens to have the convexity and monotonicity properties.



However, suppose we define

$$c^*(\mathbf{w}, y) = \min \mathbf{w}\mathbf{x}$$

such that  $\mathbf{x}$  is in  $V^*(y)$ .

What is the relationship between  $c^*(\mathbf{w}, y)$  and  $c(\mathbf{w}, y)$ ?

**Proposition 3.7.2 (Equality of  $c(\mathbf{w}, y)$  and  $c^*(\mathbf{w}, y)$ )** *It follows from the definition of the functions that  $c^*(\mathbf{w}, y) = c(\mathbf{w}, y)$ .*

PROOF. It is easy to see that  $c^*(\mathbf{w}, y) \leq c(\mathbf{w}, y)$ ; since  $V^*(y)$  always contains  $V(y)$ , the minimal cost bundle in  $V^*(y)$  must be at least as small as the minimal cost bundle in  $V(y)$ . Suppose that for some prices  $\mathbf{w}'$  the cost-minimizing bundle  $\mathbf{x}'$  in  $V^*(y)$  has the property that  $\mathbf{w}'\mathbf{x}' = c^*(\mathbf{w}', y) < c(\mathbf{w}', y)$ . But this can't happen, since by definition of  $V^*(y)$ ,  $\mathbf{w}'\mathbf{x}' \geq c(\mathbf{w}', y)$ .

This proposition shows that the cost function for the technology  $V(y)$  is the same as the cost function for its convexification  $V^*(y)$ . In this sense, the assumption of convex input requirement sets is not very restrictive from an economic point of view.

Let us summarize the discussion to date:

- (1) Given a cost function we can define an input requirement set  $V^*(y)$ .
- (2) If the original technology is convex and monotonic, the constructed technology will be identical with the original technology.
- (3) If the original technology is non-convex or monotonic, the constructed input requirement will be a convexified, monotonized version of the original set, and, most importantly, the constructed technology will have the same cost function as the original technology.

In conclusion, the cost function of a firm summarizes all of the economically relevant aspects of its technology.

**Example 3.7.1 (Recovering production from a cost function)** Suppose we are given a specific cost function  $c(\mathbf{w}, y) = yw_1^a w_2^{1-a}$ . How can we solve for its associated tech-

nology? According to the derivative property

$$\begin{aligned}x_1(\mathbf{w}, y) &= ayw_1^{a-1}w_2^{1-a} = ay \left( \frac{w_2}{w_1} \right)^{1-a} \\x_2(\mathbf{w}, y) &= (1-a)yw_1^aw_2^{-a} = (1-a)y \left( \frac{w_2}{w_1} \right)^{-a}.\end{aligned}$$

We want to eliminate  $w_2/w_1$  from these two equations and get an equation for  $y$  in terms of  $x_1$  and  $x_2$ . Rearranging each equation gives

$$\begin{aligned}\frac{w_2}{w_1} &= \left( \frac{x_1}{ay} \right)^{\frac{1}{1-a}} \\ \frac{w_2}{w_1} &= \left( \frac{x_2}{(1-a)y} \right)^{-\frac{1}{a}}.\end{aligned}$$

Setting these equal to each other and raising both sides to the  $-a(1-a)$  power,

$$\frac{x_1^{-a}}{a^{-a}y^{-a}} = \frac{x_2^{1-a}}{(1-a)^{(1-a)}y^{1-a}},$$

or,

$$[a^a(1-a)^{1-a}]y = x_1^ax_2^{1-a}.$$

This is just the Cobb-Douglas technology.

We know that if the technology exhibited constant returns to scale, then the cost function would have the form  $c(\mathbf{w})y$ . Here we show that the reverse implication is also true.

**Proposition 3.7.3 (Constant returns to scale.)** *Let  $V(y)$  be convex and monotonic; then if  $c(\mathbf{w}, y)$  can be written as  $yc(\mathbf{w})$ ,  $V(y)$  must exhibit constant returns to scale.*

PROOF Using convexity, monotonicity, and the assumed form of the cost function assumptions, we know that

$$V(y) = V^*(y) = \{\mathbf{x}: \mathbf{w} \cdot \mathbf{x} \geq yc(\mathbf{w}) \text{ for all } \mathbf{w} \geq \mathbf{0}\}.$$

We want to show that, if  $\mathbf{x}$  is in  $V^*(y)$ , then  $t\mathbf{x}$  is in  $V^*(ty)$ . If  $\mathbf{x}$  is in  $V^*(y)$ , we know that  $\mathbf{w}\mathbf{x} \geq yc(\mathbf{w})$  for all  $\mathbf{w} \geq \mathbf{0}$ . Multiplying both sides of this equation by  $t$  we get:  $\mathbf{w}t\mathbf{x} \geq tyc(\mathbf{w})$  for all  $\mathbf{w} \geq \mathbf{0}$ . But this says  $t\mathbf{x}$  is in  $V^*(ty)$ .

### 3.7.2 Characterization of Cost Functions

We have seen in the last section that all cost functions are nondecreasing, homogeneous, concave, continuous functions of prices. The question arises: suppose that you are given a nondecreasing, homogeneous, concave, continuous function of prices is it necessarily the cost function of some technology? The answer is yes, and the following proposition shows how to construct such a technology.

**Proposition 3.7.4** *Let  $\phi(\mathbf{w}, y)$  be a differentiable function satisfying*

- (1)  $\phi(t\mathbf{w}, y) = t\phi(\mathbf{w}, y)$  for all  $t \geq 0$ ;
- (2)  $\phi(\mathbf{w}, y) \geq 0$  for  $\mathbf{w} \geq 0$  and  $y \geq 0$ ;
- (3)  $\phi(\mathbf{w}', y) \geq \phi(\mathbf{w}, y)$  for  $\mathbf{w}' \geq \mathbf{w}$ ;
- (4)  $\phi(\mathbf{w}, y)$  is concave in  $\mathbf{w}$ .

*Then  $\phi(\mathbf{w}, y)$  is the cost function for the technology defined by  $V^*(y) = \{\mathbf{x} \geq \mathbf{0} : \mathbf{w}\mathbf{x} \leq \phi(\mathbf{w}, y), \text{ for all } \mathbf{w} \geq \mathbf{0}\}$ .*

PROOF. Given a  $\mathbf{w} \geq 0$  we define

$$\mathbf{x}(\mathbf{w}, y) = \left( \frac{\partial \phi(\mathbf{w}, y)}{\partial w_1}, \dots, \frac{\partial \phi(\mathbf{w}, y)}{\partial w_n} \right)$$

and note that since  $\phi(\mathbf{w}, y)$  is homogeneous of degree 1 in  $w$ , Euler's law implies that  $\phi(\mathbf{w}, y)$  can be written as

$$\phi(\mathbf{w}, y) = \sum_{i=1}^n w_i \frac{\partial \phi(\mathbf{w}, y)}{\partial w_i} = \mathbf{w}\mathbf{x}(\mathbf{w}, y).$$

Note that the monotonicity of  $\phi(\mathbf{w}, y)$  implies  $\mathbf{x}(\mathbf{w}, y) \geq 0$ .

What we need to show is that for any given  $\mathbf{w}' \geq 0$ ,  $\mathbf{x}(\mathbf{w}', y)$  actually minimizes  $\mathbf{w}'\mathbf{x}$  over all  $\mathbf{x}$  in  $V^*(y)$ :

$$\phi(\mathbf{w}', y) = \mathbf{w}'\mathbf{x}(\mathbf{w}', y) \leq \mathbf{w}'\mathbf{x} \text{ for all } \mathbf{x} \text{ in } V^*(y).$$

First, we show that  $\mathbf{x}(\mathbf{w}', y)$  is feasible; that is,  $\mathbf{x}(\mathbf{w}', y)$  is in  $V^*(y)$ . By the concavity of  $\phi(\mathbf{w}, y)$  in  $\mathbf{w}$  we have

$$\phi(\mathbf{w}', y) \leq \phi(\mathbf{w}, y) + \mathbf{D}\phi(\mathbf{w}, y)(\mathbf{w}' - \mathbf{w})$$

for all  $\mathbf{w} \geq 0$ .

Using Euler's law as above, this reduces to

$$\phi(\mathbf{w}', y) \leq \mathbf{w}' \mathbf{x}(\mathbf{w}, y) \text{ for all } \mathbf{w} \geq 0.$$

It follows from the definition of  $V^*(y)$ , that  $\mathbf{x}(\mathbf{w}', y)$  is in  $V^*(y)$ .

Next we show that  $\mathbf{x}(\mathbf{w}, y)$  actually minimizes  $\mathbf{w}\mathbf{x}$  over all  $\mathbf{x}$  in  $V^*(y)$ . If  $\mathbf{x}$  is in  $V^*(y)$ , then by definition it must satisfy

$$\mathbf{w}\mathbf{x} \geq \phi(\mathbf{w}, y).$$

But by Euler's law,

$$\phi(\mathbf{w}, y) = \mathbf{w}\mathbf{x}(\mathbf{w}, y).$$

The above two expressions imply

$$\mathbf{w}\mathbf{x} \geq \mathbf{w}\mathbf{x}(\mathbf{w}, y)$$

for all  $\mathbf{x}$  in  $V^*(y)$  as required.

### 3.7.3 The Integrability for Cost Functions

The proposition proved in the last subsection raises an interesting question. Suppose you are given a set of functions  $(g_i(\mathbf{w}, y))$  that satisfy the properties of conditional factor demand functions described in the previous sections, namely, that they are homogeneous of degree 0 in prices and that

$$\left( \frac{\partial g_i(\mathbf{w}, y)}{\partial w_j} \right)$$

is a symmetric negative semi-definite matrix. Are these functions necessarily factor demand functions for some technology?

Let us try to apply the above proposition. First, we construct a candidate for a cost function:

$$\phi(\mathbf{w}, y) = \sum_{i=1}^n w_i g_i(\mathbf{w}, y).$$

Next, we check whether it satisfies the properties required for the proposition just proved.

- 1) Is  $\phi(\mathbf{w}, y)$  homogeneous of degree 1 in  $\mathbf{w}$ ? To check this we look at  $\phi(t\mathbf{w}, y) = \sum_i tw_i g_i(t\mathbf{w}, y)$ . Since the functions  $g_i(\mathbf{w}, y)$  are by assumption homogeneous of degree 0,  $g_i(t\mathbf{w}, y) = g_i(\mathbf{w}, y)$  so that

$$\phi(t\mathbf{w}, y) = t \sum_{i=1}^n w_i g_i(\mathbf{w}, y) = t\phi(\mathbf{w}, y).$$

- 2) Is  $\phi(\mathbf{w}, y) \geq 0$  for  $\mathbf{w} \geq \mathbf{0}$ ? Since  $g_i(\mathbf{w}, y) \geq 0$ , the answer is clearly yes.

- 3) Is  $\phi(\mathbf{w}, y)$  nondecreasing in  $w_i$ ? Using the product rule, we compute

$$\frac{\partial \phi(\mathbf{w}, y)}{\partial w_i} = g_i(\mathbf{w}, y) + \sum_{j=1}^n w_j \frac{\partial g_i(\mathbf{w}, y)}{\partial w_i} = g_i(\mathbf{w}, y) + \sum_{j=1}^n w_j \frac{\partial g_i(\mathbf{w}, y)}{\partial w_j}.$$

Since  $g_i(\mathbf{w}, y)$  is homogeneous of degree 0, the last term vanishes and  $g_i(\mathbf{w}, y)$  is clearly greater than or equal to 0.

- 4) Finally is  $\phi(\mathbf{w}, y)$  concave in  $\mathbf{w}$ ? To check this we differentiate  $\phi(\mathbf{w}, y)$  twice to get

$$\left( \frac{\partial^2 \phi}{\partial w_i \partial w_j} \right) = \left( \frac{\partial g_i(\mathbf{w}, y)}{\partial w_j} \right).$$

For concavity we want these matrices to be symmetric and negative semi-definite, which they are by hypothesis.

Hence, the proposition proved in the last subsection applies and there is a technology  $V^*(y)$  that yields  $(g_i(\mathbf{w}, y))$  as its conditional factor demands. This means that the properties of homogeneity and negative semi-definiteness form a complete list of the restrictions on demand functions imposed by the model of cost-minimizing behavior.

## Reference

- Debreu, G., *Theory of Value*, (Wiley, New York), 1959.
- Diewert, E. (1974). Applications of duality theory. In M. Intriligator & D. Kendrick (Eds.), *Frontiers of Quantitative Economics*. (Amsterdam: North-Holland).
- Hicks, J. (1946). *Value and Capital*. (Oxford, England: Clarendon Press).

- Hotelling, H. (1932). Edgeworth's taxation paradox and the nature of demand and supply function. *Political Economy*, 40, 577-616.
- Jehle, G. A., and P. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, 1998, Chapter 5.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapters 2-3.
- McFadden, D. (1978). Cost, revenue, and profit functions. In M. Fuss & D. McFadden (Eds.), *Production Economics: A Dual Approach to Theory and Applications*. (Amsterdam: North-Holland).
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995, Chapter 5.
- Rubinstein Ariel, *Lecture Notes in Microeconomics (modeling the economic agent)*, Princeton University Press, 2005.
- Samuelson, P. (1947). *Foundations of Economic Analysis*. (Cambridge, Mass.: Harvard University Press).
- Shephard, R. (1953). *Cost and Production Functions*. (Princeton, NJ: Princeton University Press).
- Shephard, R. (1970). *Cost and Production Functions*. (Princeton, NJ: Princeton University Press).
- Tian, G., *Market Economics for Masses*, (with Fan Zhang), in A Series of Market Economics, Vol. 1, Ed. by G. Tian, Shanghai People's Publishing House and Hong Kong's Intelligent Book Ltd, 1993 (in Chinese), Chapters 1-2.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 1-6.

# Chapter 4

## Choice Under Uncertainty

### 4.1 Introduction

Until now, we have been concerned with the behavior of a consumer under conditions of certainty. However, many choices made by consumers take place under conditions of uncertainty. In this chapter we explore how the theory of consumer choice can be used to describe such behavior.

The board outline of this chapter parallels a standard presentation of microeconomic theory for deterministic situations. It first considers the problem of an individual consumer facing an uncertain environment. It shows how preference structures can be extended to uncertain situations and describes the nature of the consumer choice problem. We then processed to derive the expected utility theorem, a result of central importance. In the remaining sections, we discuss the concept of risk aversion, and extend the basic theory by allowing utility to depend on states of nature underlying the uncertainty as well as on the monetary payoffs. We also discuss the theory of subjective probability, which offers a way of modelling choice under uncertainty in which the probabilities of different risky alternatives are not given to the decision maker in any objective fashion.

## 4.2 Expected Utility Theory

### 4.2.1 Lotteries

The first task is to describe the set of choices facing the consumer. We shall imagine that the choices facing the consumer take the form of lotteries. Suppose there are  $S$  states. Associated with each state  $s$  is a probability  $p_s$  representing the probability that the state  $s$  will occur and a commodity bundle  $x_s$  representing the prize or reward that will be won if the state  $s$  occurs, where we have  $p_s \geq 0$  and  $\sum_{s=1}^S p_s = 1$ . The prizes may be money, bundles of goods, or even further lotteries. A lottery is denoted by

$$p_1 \circ x_1 \oplus p_2 \circ x_2 \oplus \dots \oplus p_S \circ x_S.$$

For instance, for two states, a lottery is given  $p \circ x \oplus (1 - p) \circ y$  which means: “the consumer receives prize  $x$  with probability  $p$  and prize  $y$  with probability  $(1 - p)$ .” Most situations involving behavior under risk can be put into this lottery framework. Lotteries are often represented graphically by a fan of possibilities as in Figure 4.1 below.

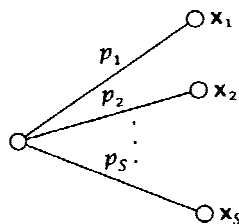


Figure 4.1: A lottery.

A compound lottery is shown in Figure 4.2. This lottery is between two prizes: a lottery between  $x$  and  $y$ , and a bundle  $z$ .

We will make several axioms about the consumers perception of the lotteries open to him.

**A1 (Certainty).**  $1 \circ x \oplus (1 - 1) \circ y \sim x$ . Getting a prize with probability one is equivalent to that prize.

**A2 (Independence of Order).**  $p \circ x \oplus (1 - p) \circ y \sim (1 - p) \circ y \oplus p \circ x$ . The consumer doesn't care about the order in which the lottery is described—only the prizes and the probabilities of winning those prizes matter.



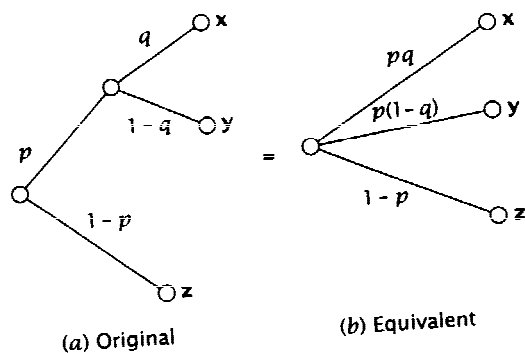


Figure 4.2: Compound lottery.

**A3 (Compounding).**  $q \circ (p \circ x \oplus (1-p) \circ y) \oplus (1-q) \circ y \sim (qp) \circ x \oplus (1-qp) \circ y$ . It is only the net probabilities of receiving the a reward that matters. It is a fundamental axiom used to reduce compound lotteries—by determining the overall probabilities associated with its components. This axiom sometimes called “reduction of compound lotteries.”

Under these assumptions we can define  $\mathcal{L}$ , the space of lotteries available to the consumer. The consumer is assumed to have preferences on this lottery space: given any two lotteries, he can choose between them. As usual we will assume the preferences are complete, reflexive, and transitive so it is an ordering preference.

The fact that lotteries have only two outcomes is not restrictive since we have allowed the outcomes to be further lotteries. This allows us to construct lotteries with arbitrary numbers of prizes by compounding two prize lotteries as shown in Figure 4.2. For example, suppose we want to represent a situation with three prizes  $x$ ,  $y$  and  $z$  where the probability of getting each prize is one third. By the reduction of compound lotteries, this lottery is equivalent to the lottery

$$\frac{2}{3} \circ \left[ \frac{1}{2} \circ x \oplus \frac{1}{2} \circ y \right] \oplus \frac{1}{3} \circ z.$$

According to the compounding axiom (A3) above, the consumer only cares about the net probabilities involved, so this is indeed equivalent to the original lottery.

## 4.2.2 Expected Utility

Under minor additional assumptions, the theorem concerning the existence of a utility function may be applied to show that there exists a continuous utility function  $u$  which

describes the consumer's preferences; that is,  $p \circ x \oplus (1 - p) \circ y \succ q \circ w \oplus (1 - q) \circ z$  if and only if

$$u(p \circ x \oplus (1 - p) \circ y) > u(q \circ w \oplus (1 - q) \circ z).$$

Of course, this utility function is not unique; any monotonic transform would do as well. Under some additional hypotheses, we can find a particular monotonic transformation of the utility function that has a very convenient property, the expected utility property:

$$u(p \circ x \oplus (1 - p) \circ y) = pu(x) + (1 - p)u(y).$$

The expected utility property says that the utility of a lottery is the expectation of the utility from its prizes and such an expected utility function is called **von Neumann-Morgenstern utility function**. To have a utility function with the above convenient property, we need the additional axioms:

**A4 (Continuity).**  $\{p \text{ in } [0, 1]: p \circ x \oplus (1 - p) \circ y \succeq z\}$  and  $\{p \text{ in } [0, 1]: z \succeq p \circ x \oplus (1 - p) \circ y\}$  are closed sets for all  $x, y$  and  $z$  in  $\mathcal{L}$ . Axiom 4 states that preferences are continuous with respect to probabilities.

**A5 (Strong Independence).**  $x \sim y$  implies  $p \circ x \oplus (1 - p) \circ z \sim p \circ y \oplus (1 - p) \circ z$ . It says that lotteries with indifferent prizes are indifferent.

In order to avoid some technical details we will make two further assumptions.

**A6 (Boundedness).** There is some best lottery  $b$  and some worst lottery  $w$ . For any  $x$  in  $\mathcal{L}$ ,  $b \succeq x \succeq w$ .

**A7 (Monotonicity).** A lottery  $p \circ b \oplus (1 - p) \circ w$  is preferred to  $q \circ b \oplus (1 - q) \circ w$  if and only if  $p > q$ .

Axiom A7 can be derived from the other axioms. It just says that if one lottery between the best prize and the worse prize is preferred to another it must be because it gives higher probability of getting the best prize.

Under these assumptions we can state the main theorem.

**Theorem 4.2.1 (Expected utility theorem)** *If  $(\mathcal{L}, \succeq)$  satisfy Axioms 1-7, there is a utility function  $u$  defined on  $\mathcal{L}$  that satisfies the expected utility property:*

$$u(p \circ x \oplus (1 - p) \circ y) = pu(x) + (1 - p)u(y)$$

PROOF. Define  $u(b) = 1$  and  $u(w) = 0$ . To find the utility of an arbitrary lottery  $z$ , set  $u(z) = p_z$  where  $p_z$  is defined by

$$p_z \circ b \oplus (1 - p_z) \circ w \sim z. \quad (4.1)$$

In this construction the consumer is indifferent between  $z$  and a gamble between the best and the worst outcomes that gives probability  $p_z$  of the best outcome.

To ensure that this is well defined, we have to check two things.

(1) Does  $p_z$  exist? The two sets  $\{p \text{ in } [0, 1]: p \circ b \oplus (1 - p) \circ w \succeq z\}$  and  $\{p \text{ in } [0, 1]: z \succeq p \circ b \oplus (1 - p) \circ w\}$  are closed and nonempty by the continuity and boundedness axioms (A4 and A6), and every point in  $[0, 1]$  is in one or the other of the two sets. Since the unit interval is connected, there must be some  $p$  in both – but this will just be the desired  $p_z$ .

(2) Is  $p_z$  unique? Suppose  $p_z$  and  $p'_z$  are two distinct numbers and that each satisfies (4.1). Then one must be larger than the other. By the monotonicity axiom A7, the lottery that gives a bigger probability of getting the best prize cannot be indifferent to one that gives a smaller probability. Hence,  $p_z$  is unique and  $u$  is well defined.

We next check that  $u$  has the expected utility property. This follows from some simple substitutions:

$$\begin{aligned} & p \circ x \oplus (1 - p) \circ y \\ & \sim_1 p \circ [p_x \circ b \oplus (1 - p_x) \circ w] \oplus (1 - p) \circ [p_y \circ b \oplus (1 - p_y) \circ w] \\ & \sim_2 [pp_x + (1 - p)p_y] \circ b \oplus [1 - pp_x - (1 - p)p_y] \circ w \\ & \sim_3 [pu(x) + (1 - p)u(y)] \circ b \oplus [1 - pu(x) - (1 - p)u(y)] \circ w. \end{aligned}$$

Substitution 1 uses the strong independence axiom (A5) and the definition of  $p_z$  and  $p_y$ . Substitution 2 uses the compounding axiom (A3), which says only the net probabilities of obtaining  $b$  or  $w$  matter. Substitution 3 uses the construction of the utility function.

It follows from the construction of the utility function that

$$u(p \circ x \oplus (1 - p) \circ y) = pu(x) + (1 - p)u(y).$$

Finally, we verify that  $u$  is a utility function. Suppose that  $x \succ y$ . Then

$$\begin{aligned} u(x) &= p_z \text{ such that } x \sim p_x \circ b \oplus (1 - p_x) \circ w \\ u(y) &= p_y \text{ such that } y \sim p_y \circ b \oplus (1 - p_y) \circ w \end{aligned}$$

By the monotonicity axiom (A7), we must have  $u(x) > u(y)$ .

### 4.2.3 Uniqueness of the Expected Utility Function

We have shown that there exists an expected utility function  $u : \mathcal{L} \rightarrow R$ . Of course, any monotonic transformation of  $u$  will also be a utility function that describes the consumer's choice behavior. But will such a monotonic transform preserve the expected utility property? Does the construction described above characterize expected utility functions in any way?

It is not hard to see that, if  $u(\cdot)$  is an expected utility function describing some consumer, then so is  $v(\cdot) = au(\cdot) + c$  where  $a > 0$ ; that is, any affine transformation of an expected utility function is also an expected utility function. This is clear since

$$\begin{aligned} v(p \circ x \oplus (1 - p) \circ y) &= au(p \circ x \oplus (1 - p) \circ y) + c \\ &= a[pu(x) + (1 - p)u(y)] + c \\ &= pv(x) + (1 - p)v(y). \end{aligned}$$

It is not much harder to see the converse: that any monotonic transform of  $u$  that has the expected utility property must be an affine transform. Stated another way:

**Theorem 4.2.2 ( Uniqueness of expected utility function )** *An expected utility function is unique up to an affine transformation.*

PROOF. According to the above remarks we only have to show that, if a monotonic transformation preserves the expected utility property, it must be an affine transformation. Let  $f : R \rightarrow R$  be a monotonic transformation of  $u$  that has the expected utility property. Then

$$f(u(p \circ x \oplus (1 - p) \circ y)) = pf(u(x)) + (1 - p)f(u(y)),$$

or

$$f(pu(x) + (1 - p)u(y)) = pf(u(x)) + (1 - p)f(u(y)).$$

But this is equivalent to the definition of an affine transformation.

#### 4.2.4 Other Notations for Expected Utility

We have proved the expected utility theorem for the case where there are two outcomes to the lotteries. As indicated earlier, it is straightforward to extend this proof to the case of a finite number of outcomes by using compound lotteries. If outcome  $x_i$  is received with probability  $p_i$  for  $i = 1, \dots, n$ , the expected utility of this lottery is simply

$$\sum_{i=1}^n p_i u(x_i). \quad (4.2)$$

Subject to some minor technical details, the expected utility theorem also holds for continuous probability distributions. If  $p(x)$  is probability density function defined on outcomes  $x$ , then the expected utility of this gamble can be written as

$$\int u(x)p(x)dx. \quad (4.3)$$

We can subsume of these cases by using the expectation operator. Let  $X$  be a random variable that takes on values denoted by  $x$ . Then the utility function of  $X$  is also a random variable,  $u(X)$ . The expectation of this random variable  $Eu(X)$  is simply the expected utility associated with the lottery  $X$ . In the case of a discrete random variable,  $Eu(X)$  is given by (4.2), and in the case of a continuous random variable  $Eu(X)$  is given by (4.3).

### 4.3 Risk aversion

#### 4.3.1 Absolute Risk Aversion

Let us consider the case where the lottery space consists solely of gambles with money prizes. We have shown that if the consumer's choice behavior satisfies Axioms 1-7, we can use an expected utility function to represent the consumer's preferences on lotteries. This means that we can describe the consumer's behavior over all money gambles by expected utility function. For example, to compute the consumer's expected utility of a gamble  $p \circ x \oplus (1 - p) \circ y$ , we just look at  $pu(x) + (1 - p)u(y)$ .

This construction is illustrated in Figure 4.3 for  $p = \frac{1}{2}$ . Notice that in this example the consumer prefers to get the expected value of the lottery. This is, the utility of the lottery  $u(p \circ x \oplus (1 - p) \circ y)$  is less than the utility of the expected value of the lottery,

$px + (1 - p)y$ . Such behavior is called risk aversion. A consumer may also be risk loving; in such a case, the consumer prefers a lottery to its expected value.

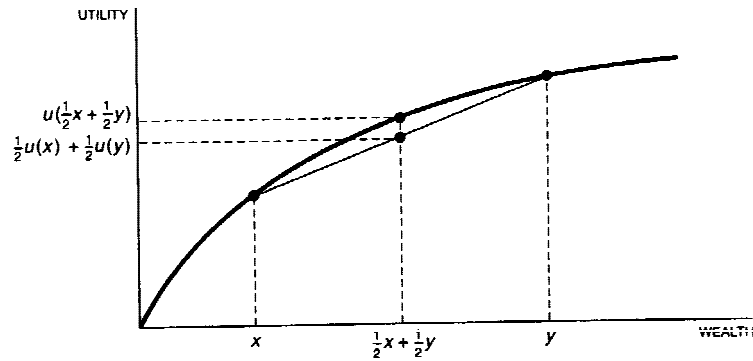


Figure 4.3: Expected utility of a gamble.

If a consumer is risk averse over some region, the chord drawn between any two points of the graph of his utility function in this region must lie below the function. This is equivalent to the mathematical definition of a concave function. Hence, **concavity of the expected utility function is equivalent to risk aversion.**

It is often convenient to have a measure of risk aversion. Intuitively, the more concave the expected utility function, the more risk averse the consumer. Thus, we might think we could measure risk aversion by the second derivative of the expected utility function. However, this definition is not invariant to changes in the expected utility function: if we multiply the expected utility function by 2, the consumer's behavior doesn't change, but our proposed measure of risk aversion does. However, if we normalize the second derivative by dividing by the first, we get a reasonable measure, known as the Arrow-Pratt measure of (absolute) risk aversion:

$$r(w) = -\frac{u''(w)}{u'(w)}$$

**Example 4.3.1 (Constant risk aversion)** Suppose an individual has a constant risk aversion coefficient  $r$ . Then the utility function satisfies

$$u''(x) = -ru'(x).$$

One can easily check that all solutions are

$$u(x) = -ae^{-rx} + b$$

where  $a$  and  $b$  are arbitrary. For  $u(x)$  to be increasing in  $x$ , we must take  $a > 0$ .

**Example 4.3.2** A woman with current wealth  $X$  has the opportunity to bet any amount on the occurrence of an event that she knows will occur with probability  $p$ . If she wagers  $w$ , she will receive  $2w$  if the event occurs and 0 if it does not. She has a constant risk aversion coefficient utility  $u(x) = -e^{-rx}$  with  $r > 0$ . How much should she wager?

Her final wealth will be either  $X + w$  or  $X - w$ . Hence she solves

$$\max_w \{pu(X + w) + (1 - p)u(X - w)\} = \max_w \{-pe^{-r(X+w)} - (1 - p)e^{-r(X-w)}\}$$

Setting the derivative to zero yields

$$(1 - p)e^{rw} = pe^{-rw}.$$

Hence,

$$w = \frac{1}{2r} \ln \frac{p}{1 - p}.$$

Note that a positive wager will be made for  $p > 1/2$ . The wager decreases as the risk coefficient increases. Note also that in this case the results is independent of the initial wealth—a particular feature of this utility function.

**Example 4.3.3 (The demand for insurance)** Suppose a consumer initially has monetary wealth  $W$ . There is some probability  $p$  that he will lose an amount  $L$  for example, there is some probability his house will burn down. The consumer can purchase insurance that will pay him  $q$  dollars in the event that he incurs this loss. The amount of money that he has to pay for  $q$  dollars of insurance coverage is  $\pi q$ ; here  $\pi$  is the premium per dollar of coverage.

How much coverage will the consumer purchase? We look at the utility maximization problem

$$\max q u(W - L - \pi q + q) + (1 - p)u(W - \pi q).$$

Taking the derivative with respect to  $q$  and setting it equal to zero, we find

$$pu'(W - L + q^*(1 - \pi))(1 - \pi) - (1 - p)u'(W - \pi q^*)\pi = 0$$

$$\frac{u'(W - L + (1 - \pi)q^*)}{u'(W - \pi q^*)} = \frac{(1 - p)}{p} \frac{\pi}{1 - \pi}$$

If the event occurs, the insurance company receives  $\pi q - q$  dollars. If the event doesn't occur, the insurance company receives  $\pi q$  dollars. Hence, the expected profit of the company is

$$(1 - p)\pi q - p(1 - \pi)q.$$

Let us suppose that competition in the insurance industry forces these profits to zero. This means that

$$-p(1 - \pi)q + (1 - p)\pi q = 0,$$

from which it follows that  $\pi = p$ .

Under the zero-profit assumption the insurance firm charges an actuarially fair premium: the cost of a policy is precisely its expected value, so that  $p = \pi$ . Inserting this into the first-order conditions for utility maximization, we find

$$u'(W - L + (1 - \pi)q^*) = u'(W - \pi q^*).$$

If the consumer is strictly risk averse so that  $u''(W) < 0$ , then the above equation implies

$$W - L + (1 - \pi)q^* = W - \pi q^*$$

from which it follows that  $L = q^*$ . Thus, the consumer will completely insure himself against the loss  $L$ .

This result depends crucially on the assumption that the consumer cannot influence the probability of loss. If the consumer's actions do affect the probability of loss, the insurance firms may only want to offer partial insurance, so that the consumer will still have an incentive to be careful.

### 4.3.2 Global Risk Aversion

The Arrow-Pratt measure seems to be a sensible interpretation of local risk aversion: one agent is more risk averse than another if he is willing to accept fewer small gambles. However, in many circumstances we want a global measure of risk aversion—that is, we want to say that one agent is more risk averse than another for all levels of wealth. What are natural ways to express this condition?



The first plausible way is to formalize the notion that an agent with utility function  $A(w)$  is more risk averse than an agent with utility function  $B(w)$  is to require that

$$-\frac{A''(w)}{A'(w)} > -\frac{B''(w)}{B'(w)}$$

for all levels of wealth  $w$ . This simply means that agent  $A$  has a higher degree of risk aversion than agent  $B$  everywhere.

Another sensible way to formalize the notion that agent  $A$  is more risk averse than agent  $B$  is to say that agent  $A$ 's utility function is “more concave” than agent  $B$ 's. More precisely, we say that agent  $A$ 's utility function is a concave transformation of agent  $B$ 's; that is, there exists some increasing, strictly concave function  $G(\cdot)$  such that

$$A(w) = G(B(w)).$$

A third way to capture the idea that  $A$  is more risk averse than  $B$  is to say that  $A$  would be willing to pay more to avoid a given risk than  $B$  would. In order to formalize this idea, let  $\tilde{\epsilon}$  be a random variable with expectation of zero:  $E\tilde{\epsilon} = 0$ . Then define  $\pi_A(\tilde{\epsilon})$  to be the maximum amount of wealth that person  $A$  would give up in order to avoid facing the random variable  $\tilde{\epsilon}$ . In symbols, this risk premium is

$$A(w - \pi_A(\tilde{\epsilon})) = EA(w + \tilde{\epsilon}).$$

The left-hand side of this expression is the utility from having wealth reduced by  $\pi_A(\tilde{\epsilon})$  and the right-hand side is the expected utility from facing the gamble  $\tilde{\epsilon}$ . It is natural to say that person  $A$  is (globally) more risk averse than person  $B$  if  $\pi_A(\tilde{\epsilon}) > \pi_B(\tilde{\epsilon})$  for all  $\epsilon$  and  $w$ .

It may seem difficult to choose among these three plausible sounding interpretations of what it might mean for one agent to be “globally more risk averse” than another. Luckily, it is not necessary to do so: all three definitions turn out to be equivalent! As one step in the demonstration of this fact we need the following result, which is of great use in dealing with expected utility functions.

**Lemma 4.3.1 ( Jensen's inequality )** *Let  $X$  be a nondegenerate random variable and  $f(X)$  be a strictly concave function of this random variable. Then  $Ef(X) < f(EX)$ .*

PROOF. This is true in general, but is easiest to prove in the case of a differentiable concave function. Such a function has the property that at any point  $\bar{x}$ ,  $f(x) < f(\bar{x}) + f'(\bar{x})(x - \bar{x})$ . Let  $\bar{X}$  be the expected value of  $X$  and take expectations of each side of this expression, we have

$$Ef(X) < f(\bar{X}) + f'(\bar{X})E(X - \bar{X}) = f(\bar{X}),$$

from which it follows that

$$Ef(X) < f(\bar{X}) = f(EX).$$

**Theorem 4.3.1 (Pratt's theorem)** *Let  $A(w)$  and  $B(w)$  be two differentiable, increasing and concave expected utility functions of wealth. Then the following properties are equivalent.*

1.  $-A''(w)/A'(w) > -B''(w)/B'(w)$  for all  $w$ .
2.  $A(w) = G(B(w))$  for some increasing strictly concave function  $G$ .
3.  $\pi_A(\tilde{\epsilon}) > \pi_B(\tilde{\epsilon})$  for all random variables  $\tilde{\epsilon}$  with  $E\tilde{\epsilon} = 0$ .

PROOF

(1) implies (2). Define  $G(B)$  implicitly by  $A(w) = G(B(w))$ . Note that monotonicity of the utility functions implies that  $G$  is well defined i.e., that there is a unique value of  $G(B)$  for each value of  $B$ . Now differentiate this definition twice to find

$$\begin{aligned} A'(w) &= G'(B)B'(w) \\ A''(w) &= G''(B)B'(w)^2 + G'(B)B''(w). \end{aligned}$$

Since  $A'(w) > 0$  and  $B'(w) > 0$ , the first equation establishes  $G'(B) > 0$ . Dividing the second equation by the first gives us

$$\frac{A''(w)}{A'(w)} = \frac{G''(B)}{G'(B)}B'(w) + \frac{B''(w)}{B'(w)}.$$

Rearranging gives us

$$\frac{G''(B)}{G'(B)}B'(w) = \frac{A''(w)}{A'(w)} - \frac{B''(w)}{B'(w)} < 0,$$

where the inequality follows from (1). This shows that  $G''(B) < 0$ , as required.

(2) implies (3). This follows from the following chain of inequalities:

$$\begin{aligned}
A(w - \pi_A) &= EA(w + \tilde{\epsilon}) = EG(B(w + \tilde{\epsilon})) \\
&< G(EB(w + \tilde{\epsilon})) = G(B(w - \pi_B)) \\
&= A(w - \pi_B).
\end{aligned}$$

All of these relationships follow from the definition of the risk premium except for the inequality, which follows from Jensen's inequality. Comparing the first and the last terms, we see that  $\pi_A > \pi_B$ .

(3) implies (1). Since (3) holds for all zero-mean random variables  $\tilde{\epsilon}$ , it must hold for arbitrarily small random variables. Fix an  $\tilde{\epsilon}$ , and consider the family of random variables defined by  $t\tilde{\epsilon}$  for  $t$  in  $[0, 1]$ . Let  $\pi(t)$  be the risk premium as a function of  $t$ . The second-order Taylor series expansion of  $\pi(t)$  around  $t = 0$  is given by

$$\pi(t) \approx \pi(0) + \pi'(0)t + \frac{1}{2}\pi''(0)t^2. \quad (4.4)$$

We will calculate the terms in this Taylor series in order to see how  $\pi(t)$  behaves for small  $t$ . The definition of  $\pi(t)$  is

$$A(w - \pi(t)) \equiv EA(w + t\tilde{\epsilon}).$$

It follows from this definition that  $\pi(0) = 0$ . Differentiating the definition twice with respect to  $t$  gives us

$$\begin{aligned}
-A'(w - \pi(t))\pi'(t) &= E[A'(w + t\tilde{\epsilon})\tilde{\epsilon}] \\
A''(w - \pi(t))\pi'(t)^2 - A'(w - \pi(t))\pi''(t) &= E[A''(w + t\tilde{\epsilon})\tilde{\epsilon}].
\end{aligned}$$

Evaluating the first expression when  $t = 0$ , we see that  $\pi'(0) = 0$ . Evaluating the second expression when  $t = 0$ , we see that

$$\pi''(0) = -\frac{EA''(w)\tilde{\epsilon}^2}{A'(w)} = -\frac{A''(w)}{A'(w)}\sigma^2,$$

where  $\sigma^2$  is the variance of  $\tilde{\epsilon}$ . Plugging the derivatives into equation (4.4) for  $\pi(t)$ , we have

$$\pi(t) \approx 0 + 0 - \frac{A''(w)}{A'(w)} \frac{\sigma^2}{2} t^2.$$

This implies that for arbitrarily small values of  $t$ , the risk premium depends monotonically on the degree of risk aversion, which is what we wanted to show.

### 4.3.3 Relative Risk Aversion

Consider a consumer with wealth  $w$  and suppose that she is offered gambles of the form: with probability  $p$  she will receive  $x$  percent of her current wealth; with probability  $(1 - p)$  she will receive  $y$  percent of her current wealth. If the consumer evaluates lotteries using expected utility, the utility of this lottery will be

$$pu(xw) + (1 - p)u(yw).$$

Note that this multiplicative gamble has a different structure than the additive gambles analyzed above. Nevertheless, relative gambles of this sort often arise in economic problems. For example, the return on investments is usually stated relative to the level of investment.

Just as before we can ask when one consumer will accept more small relative gambles than another at a given wealth level. Going through the same sort of analysis used above, we find that the appropriate measure turns out to be the Arrow-Pratt measure of relative risk aversion:

$$\rho = -\frac{u''(w)w}{u'(w)}.$$

It is reasonable to ask how absolute and relative risk aversions might vary with wealth. It is quite plausible to assume that absolute risk aversion decreases with wealth: as you become more wealthy you would be willing to accept more gambles expressed in absolute dollars. The behavior of relative risk aversion is more problematic; as your wealth increases would you be more or less willing to risk losing a specific fraction of it? Assuming constant relative risk aversion is probably not too bad an assumption, at least for small changes in wealth.

**Example 4.3.4 (Mean-variance utility)** In general the expected utility of a gamble depends on the entire probability distribution of the outcomes. However, in some circumstances the expected utility of a gamble will only depend on certain summary statistics of the distribution. The most common example of this is a mean-variance utility function.

For example, suppose that the expected utility function is quadratic, so that  $u(w) = w - bw^2$ . Then expected utility is

$$Eu(w) = Ew - bEw^2 = \bar{w} - b\bar{w}^2 - b\sigma_w^2.$$

Hence, the expected utility of a gamble is only a function of the mean and variance of wealth.

Unfortunately, the quadratic utility function has some undesirable properties: it is a decreasing function of wealth in some ranges, and it exhibits increasing absolute risk aversion.

A more useful case when mean-variance analysis is justified is the case when wealth is Normally distributed. It is well-known that the mean and variance completely characterize a Normal random variable; hence, choice among Normally distributed random variables reduces to a comparison on their means and variances. One particular case that is of special interest is when the consumer has a utility function of the form  $u(w) = -e^{-rw}$  which exhibits constant absolute risk aversion. Furthermore, when wealth is Normally distributed

$$Eu(w) = - \int e^{-rw} f z(w) dw = -e^{-r[\bar{w} - r\sigma_w^2/2]}.$$

(To do the integration, either complete the square or else note that this is essentially the calculation that one does to find the moment generating function for the Normal distribution.) Note that expected utility is increasing in  $\bar{w} - r\sigma_w^2/2$ . This means that we can take a monotonic transformation of expected utility and evaluate distributions of wealth using the utility function  $u(\bar{w}, \sigma_w^2) = \bar{w} - \frac{r}{2}\sigma_w^2$ . This utility function has the convenient property that it is linear in the mean and variance of wealth.

## 4.4 State Dependent Utility

In our original analysis of choice under uncertainty, the prizes were simply abstract bundles of goods; later we specialized to lotteries with only monetary outcomes when considering the risk aversion issue. However, this is restrictive. After all, a complete description of the outcome of a dollar gamble should include not only the amount of money available in each outcome but also the prevailing prices in each outcome.

More generally, the usefulness of a good often depends on the circumstances or state of nature in which it becomes available. An umbrella when it is raining may appear very different to a consumer than an umbrella when it is not raining. These examples show that in some choice problems it is important to distinguish goods by the state of nature

in which they are available.

For example, suppose that there are two states of nature, hot and cold, which we index by  $h$  and  $c$ . Let  $x_h$  be the amount of ice cream delivered when it is hot and  $x_c$  the amount delivered when it is cold. Then if the probability of hot weather is  $p$ , we may write a particular lottery as  $pu(h, x_h) + (1 - p)u(c, x_c)$ . Here the bundle of goods that is delivered in one state is “hot weather and  $x_h$  units of ice cream,” and “cold weather and  $x_c$ , units of ice cream” in the other state.

A more serious example involves health insurance. The value of a dollar may well depend on one’s health – how much would a million dollars be worth to you if you were in a coma? In this case we might well write the utility function as  $u(h, m_h)$  where  $h$  is an indicator of health and  $m$  is some amount of money. These are all examples of state-dependent utility functions. This simply means that the preferences among the goods under consideration depend on the state of nature under which they become available.

## 4.5 Subjective Probability Theory

In the discussion of expected utility theory we have used “objective” probabilities — such as probabilities calculated on the basis of some observed frequencies and asked what axioms about a person’s choice behavior would imply the existence of an expected utility function that would represent that behavior.

However, many interesting choice problems involve subjective probabilities: a given agent’s perception of the likelihood of some event occurring. Similarly, we can ask what axioms about a person’s choice behavior can be used to infer the existence of subjective probabilities; i.e., that the person’s choice behavior can be viewed as if he were evaluating gambles according to their expected utility with respect to some subjective probability measures.

As it happens, such sets of axioms exist and are reasonably plausible. Subjective probabilities can be constructed in a way similar to the manner with which the expected utility function was constructed. Recall that the utility of some gamble  $x$  was chosen to be that number  $u(x)$  such that

$$x \sim u(x) \circ b \oplus (1 - u(x)) \circ w.$$

Suppose that we are trying to ascertain an individual's subjective probability that it will rain on a certain date. Then we can ask at what probability  $p$  will the individual be indifferent between the gamble  $p \circ b \oplus (1 - p) \circ w$  and the gamble "Receive  $b$  if it rains and  $w$  otherwise."

More formally, let  $E$  be some event, and let  $p(E)$  stand for the (subjective) probability that  $E$  will occur. We define the subjective probability that  $E$  occurs by the number  $p(E)$  that satisfies  $p(E) \circ b \oplus (1 - p(E)) \circ w \sim$  receive  $b$  if  $E$  occurs and  $w$  otherwise.

It can be shown that under certain regularity assumptions the probabilities defined in this way have all of the properties of ordinary objective probabilities. In particular, they obey the usual rules for manipulation of conditional probabilities. This has a number of useful implications for economic behavior. We will briefly explore one such implication. Suppose that  $p(H)$  is an individual's subjective probability that a particular hypothesis is true, and that  $E$  is an event that is offered as evidence that  $H$  is true. How should a rational economic agent adjust his probability belief about  $H$  in light of the evidence  $E$ ? That is, what is the probability of  $H$  being true, conditional on observing the evidence  $E$ ?

We can write the joint probability of observing  $E$  and  $H$  being true as

$$p(H, E) = p(H|E)p(E) = p(E|H)p(H).$$

Rearranging the right-hand sides of this equation,

$$p(H|E) = \frac{p(E|H)p(H)}{p(E)}.$$

This is a form of Bayes' law which relates the prior probability  $p(H)$ , the probability that the hypothesis is true before observing the evidence, to the posterior probability, the probability that the hypothesis is true after observing the evidence.

Bayes' law follows directly from simple manipulations of conditional probabilities. If an individual's behavior satisfies restrictions sufficient to ensure the existence of subjective probabilities, those probabilities must satisfy Bayes' law. Bayes' law is important since it shows how a rational individual should update his probabilities in the light of evidence, and hence serves as the basis for most models of rational learning behavior.

Thus, both the utility function and the subjective probabilities can be constructed from observed choice behavior, as long as the observed choice behavior follows certain

intuitively plausible axioms. However, it should be emphasized that although the axioms are intuitively plausible it does not follow that they are accurate descriptions of how individuals actually behave. That determination must be based on empirical evidence. Expected utility theory and subjective probability theory were motivated by considerations of rationality. The axioms underlying expected utility theory seem plausible, as does the construction that we used for subjective probabilities.

Unfortunately, real-life individual behavior appears to systematically violate some of the axioms. Here we present two famous examples.

**Example 4.5.1 (The Allais paradox)** You are asked to choose between the following two gambles:

**Gamble A.** A 100 percent chance of receiving 1 million.

**Gamble B.** A 10 percent chance of 5 million, an 89 percent chance of 1 million, and a 1 percent chance of nothing.

Before you read any further pick one of these gambles, and write it down. Now consider the following two gambles.

**Gamble C.** An 11 percent chance of 1 million, and an 89 percent chance of nothing.

**Gamble D.** A 10 percent chance of 5 million, and a 90 percent chance of nothing.

Again, please pick one of these two gambles as your preferred choice and write it down. Many people prefer A to B and D to C. However, these choices violate the expected utility axioms! To see this, simply write the expected utility relationship implied by  $A \succeq B$ :

$$u(1) > .1u(5) + .89u(1) + .01u(0).$$

Rearranging this expression gives

$$.11u(1) > .1u(5) + .01u(0),$$

and adding  $.89u(0)$  to each side yields

$$.11u(1) + .89u(0) > .1u(5) + .90u(0),$$

It follows that gamble C must be preferred to gamble D by an expected utility maximizer.



**Example 4.5.2 (The Ellsberg paradox)** The Ellsberg paradox concerns subjective probability theory. You are told that an urn contains 300 balls. One hundred of the balls are red and 200 are either blue or green.

**Gamble A.** You receive \$1,000 if the ball is red.

**Gamble B.** You receive \$1,000 if the ball is blue.

Write down which of these two gambles you prefer. Now consider the following two gambles:

**Gamble C.** You receive \$1,000 if the ball is not red.

**Gamble D.** You receive \$1,000 if the ball is not blue.

It is common for people to strictly prefer A to B and C to D. But these preferences violate standard subjective probability theory. To see why, let  $R$  be the event that the ball is red, and  $\neg R$  be the event that the ball is not red, and define  $B$  and  $\neg B$  accordingly. By ordinary rules of probability,

$$p(R) = 1 - p(\neg R) \quad (4.5)$$

$$p(B) = 1 - p(\neg B).$$

Normalize  $u(0) = 0$  for convenience. Then if A is preferred to B, we must have  $p(R)u(1000) > p(B)u(1000)$ , from which it follows that

$$p(R) > p(B). \quad (4.6)$$

If C is preferred to D, we must have  $p(\neg R)u(1000) > p(\neg B)u(1000)$ , from which it follows that

$$p(\neg R) > p(\neg B). \quad (4.7)$$

However, it is clear that expressions (4.5), (4.6), and (4.7) are inconsistent.

The Ellsberg paradox seems to be due to the fact that people think that betting for or against  $R$  is “safer” than betting for or against “blue.”

Opinions differ about the importance of the Allais paradox and the Ellsberg paradox. Some economists think that these anomalies require new models to describe people’s behavior. Others think that these paradoxes are akin to “optical illusions.” Even though people are poor at judging distances under some circumstances doesn’t mean that we need to invent a new concept of distance.

## Reference

- Anscombe, F. & Aumann, R. (1963). A definition of subjective probability. *Annals of Mathematical Statistics*, 3, 199-205.
- Arrow, K. (1970). *Essays in the Theory of Risk Bearing*. (Chicago: Markham).
- Jehle, G. A., and P. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, 1998, Chapter 4.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 11.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995, Chapter 6.
- Neumann, J. & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. (Princeton, NJ: Princeton University Press).
- Pratt, J. (1964). Risk aversion in the small and in the large. *Econometrica*, 32, 122-136.
- Rubinstein Ariel, *Lecture Notes in Microeconomics (modeling the economic agent)*, Princeton University Press, 2005.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 1-6.
- Yaari, M. (1969). Some remarks on measures of risk aversion and their uses. *Journal of Economic Theory*, 1, 315-329.

## Part II

# Strategic Behavior and Markets

We have restricted ourselves until now to the ideal situation (benchmark case) where the behavior of the others are summarized in non-individualized parameters – the prices of commodities, each individual makes decision independently by taking prices as given and individuals' behavior are indirectly interacted through prices. This is clearly a very restricted assumption. In many cases, one individual's action can directly affect the actions of the others, and also are affected by the others's actions. Thus, it is very important to study the realistic case where individuals' behavior are affected each other and they interact each other. The game theory is thus a powerful tool to study individuals' cooperation and solving possible conflicts.

In this part, we will first discuss the game theory and then the various types of market structures where one individual's decision may affect the decisions of the others. It is greatly acknowledged that the materials are largely drawn from the references in the end of each chapter, especially from Varian (1992).

# Chapter 5

## Game Theory

### 5.1 Introduction

Game theory is the study of interacting decision makers. In earlier chapters we studied the theory of optimal decision making by a single agent—firm or a consumer—in very simple environments. The strategic interactions of the agents were not very complicated. In this chapter we will lay the foundations for a deeper analysis of the behavior of economic agents in more complex environments.

There are many directions from which one could study interacting decision makers. One could examine behavior from the viewpoint of sociology, psychology, biology, etc. Each of these approaches is useful in certain contexts. Game theory emphasizes a study of cold-blooded “rational” decision making, since this is felt to be the most appropriate model for most economic behavior.

Game theory has been widely used in economics in the last two decades, and much progress has been made in clarifying the nature of strategic interaction in economic models. Indeed, most economic behavior can be viewed as special cases of game theory, and a sound understanding of game theory is a necessary component of any economist’s set of analytical tools.

## 5.2 Description of a game

There are several ways of describing a game. For our purposes, the strategic form and the extensive form will be sufficient. Roughly speaking the extensive form provides an “extended” description of a game while the strategic form provides a “reduced” summary of a game. We will first describe the strategic form, reserving the discussion of the extensive form for the section on sequential games.

### 5.2.1 Strategic Form

The strategic form of the game is defined by exhibiting a set of players  $N = \{1, 2, \dots, n\}$ . Each player  $i$  has a set of strategies  $S_i$  from which he/she can choose an action  $s_i \in S_i$  and a payoff function,  $\phi_i(s)$ , that indicate the utility that each player receives if a particular combination  $s$  of strategies is chosen, where  $s = (s_1, s_2, \dots, s_n) \in S = \prod_{i=1}^n S_i$ . For purposes of exposition, we will treat two-person games in this chapter. All of the concepts described below can be easily extended to multi-person contexts.

We assume that the description of the game – the payoffs and the strategies available to the players – are common knowledge. That is, each player knows his own payoffs and strategies, and the other player’s payoffs and strategies. Furthermore, each player knows that the other player knows this, and so on. We also assume that it is common knowledge that each player is “fully rational.” That is, each player can choose an action that maximizes his utility given his subjective beliefs, and that those beliefs are modified when new information arrives according to Bayes’ law.

Game theory, by this account, is a generalization of standard, one-person decision theory. How should a rational expected utility maximizer behave in a situation in which his payoff depends on the choices of another rational expected utility maximizer? Obviously, each player will have to consider the problem faced by the other player in order to make a sensible choice. We examine the outcome of this sort of consideration below.

**Example 5.2.1 (Matching pennies)** In this game, there are two players, Row and Column. Each player has a coin which he can arrange so that either the head side or the tail side is face-up. Thus, each player has two strategies which we abbreviate as Heads or Tails. Once the strategies are chosen there are payoffs to each player which depend on

the choices that both players make.

These choices are made independently, and neither player knows the other's choice when he makes his own choice. We suppose that if both players show heads or both show tails, then Row wins a dollar and Column loses a dollar. If, on the other hand, one player exhibits heads and the other exhibits tails, then Column wins a dollar and Row loses a dollar.

|     |       | Column    |           |
|-----|-------|-----------|-----------|
|     |       | Heads     | Tails     |
| Row | Heads | $(1, -1)$ | $(-1, 1)$ |
|     | Tails | $(-1, 1)$ | $(1, -1)$ |

Table 5.1: Game Matrix of Matching Pennies

We can depict the strategic interactions in a game matrix. The entry in box (Head, Tails) indicates that player Row gets  $-1$  and player Column gets  $+1$  if this particular combination of strategies is chosen. Note that in each entry of this box, the payoff to player Row is just the negative of the payoff to player Column. In other words, this is a zero-sum game. In zero-sum games the interests of the players are diametrically opposed and are particularly simple to analyze. However, most games of interest to economists are not zero sum games.

**Example 5.2.2 (The Prisoners Dilemma)** Again we have two players, Row and Column, but now their interests are only partially in conflict. There are two strategies: to Cooperate or to Defect.

In the original story, Row and Column were two prisoners who jointly participated in a crime. They could cooperate with each other and refuse to give evidence (i.e., do not confess), or one could defect (i.e, confess) and implicate the other. They are held in separate cells, and each is privately told that if he is the only one to confess, then he will be rewarded with a light sentence of 1 year while the recalcitrant prisoner will go to jail for 10 years. However, if the person is the only one not to confess, then it is the who will serve the 10-year sentence. If both confess, they will both be shown some mercy: they will each get 5 years. Finally, if neither confesses, it will still possible to convict both of

a lesser crime that carries a sentence of 2 years. Each player wishes to minimize the time he spends in jail. The outcome can be shown in Table 5.2.

|            |               | Prisoner 2    |             |
|------------|---------------|---------------|-------------|
|            |               | Don't Confess | Confess     |
| Prisoner 1 | Don't Confess | $(-2, -2)$    | $(-10, -1)$ |
|            | Don't confess | $(-1, -10)$   | $(-5, -5)$  |

Table 5.2: The Prisoner's Dilemma

The problem is that each party has an incentive to confess, regardless of what he or she believes the other party will do. In this prisoner's dilemma, "confession" is the best strategy to each prisoner regardless the choice of the other.

An especially simple revised version of the above prisoner's dilemma given by Aumann (1987) is the game in which each player can simply announce to a referee: "Give me \$1,000," or "Give the other player \$3,000." Note that the monetary payments come from a third party, not from either of the players; the Prisoner's Dilemma is a variable-sum game.

The players can discuss the game in advance but the actual decisions must be independent. The Cooperate strategy is for each person to announce the \$3,000 gift, while the Defect strategy is to take the \$1,000 (and run!). Table 5.3 depicts the payoff matrix to the Aumann version of the Prisoner's Dilemma, where the units of the payoff are thousands of dollars.

|     |           | Column    |          |
|-----|-----------|-----------|----------|
|     |           | Cooperate | Defect   |
| Row | Cooperate | $(3, 3)$  | $(0, 4)$ |
|     | Defect    | $(4, 0)$  | $(1, 1)$ |

Table 5.3: A Revised Version of Prisoner's Dilemma by Aumann

We will discuss this game in more detail below. Again, each party has an incentive to defect, regardless of what he or she believes the other party will do. For if I believe that the other person will cooperate and give me a \$3,000 gift, then I will get \$4,000 in total



by defecting. On the other hand, if I believe that the other person will defect and just take the \$1,000, then I do better by taking \$1,000 for myself.

In other applications, cooperate and defect could have different meanings. For example, in a duopoly situation, cooperate could mean “keep charging a high price” and defect could mean “cut your price and steal your competitor’s market.”

**Example 5.2.3 (Cournot Duopoly)** Consider a simple duopoly game, first analyzed by Cournot (1838). We suppose that there are two firms who produce an identical good with a marginal cost  $c$ . Each firm must decide how much output to produce without knowing the production decision of the other duopolist. If the firms produce a total of  $x$  units of the good, the market price will be  $p(x)$ ; that is,  $p(x)$  is the inverse demand curve facing these two producers.

If  $x_i$  is the production level of firm  $i$ , the market price will then be  $p(x_1 + x_2)$ , and the profits of firm  $i$  are given by  $\pi_i(p(x_1 + x_2) - c)x_i$ . In this game the strategy of firm  $i$  is its choice of production level and the payoff to firm  $i$  is its profits.

**Example 5.2.4 (Bertrand duopoly)** Consider the same setup as in the Cournot game, but now suppose that the strategy of each player is to announce the price at which he would be willing to supply an arbitrary amount of the good in question. In this case the payoff function takes a radically different form. It is plausible to suppose that the consumers will only purchase from the firm with the lowest price, and that they will split evenly between the two firms if they charge the same price. Letting  $x(p)$  represent the market demand function and  $c$  the marginal cost, this leads to a payoff to firm 1 of the form:

$$\pi_1(p_1, p_2) = \begin{cases} (p_1 - c)x(p_1) & \text{if } p_1 < p_2 \\ (p_1 - c)x(p_1)/2 & \text{if } p_1 = p_2 \\ 0 & \text{if } p_1 > p_2 \end{cases}.$$

This game has a similar structure to that of the Prisoner’s Dilemma. If both players cooperate, they can charge the monopoly price and each reap half of the monopoly profits. But the temptation is always there for one player to cut its price slightly and thereby capture the entire market for itself. But if both players cut price, then they are both worse off.

Note that the Cournot game and the Bertrand game have a radically different structure, even though they purport to model the same economic phenomena – a duopoly. In the Cournot game, the payoff to each firm is a continuous function of its strategic choice; in the Bertrand game, the payoffs are discontinuous functions of the strategies. As might be expected, this leads to quite different equilibria. Which of these models is reasonable? The answer is that it depends on what you are trying to model. In most economic modelling, there is an art to choosing a representation of the strategy choices of the game that captures an element of the real strategic iterations, while at the same time leaving the game simple enough to analyze.

## 5.3 Solution Concepts

### 5.3.1 Mixed Strategies and Pure Strategies

In many games the nature of the strategic interaction suggests that a player wants to choose a strategy that is not predictable in advance by the other player. Consider, for example, the Matching Pennies game described above. Here it is clear that neither player wants the other player to be able to predict his choice accurately. Thus, it is natural to consider a random strategy of playing heads with some probability  $p_h$  and tails with some probability  $p_t$ . Such a strategy is called a **mixed strategy**. Strategies in which some choice is made with probability 1 are called **pure strategies**.

If  $R$  is the set of pure strategies available to Row, the set of mixed strategies open to Row will be the set of all probability distributions over  $R$ , where the probability of playing strategy  $r$  in  $R$  is  $p_r$ . Similarly,  $p_c$ , will be the probability that Column plays some strategy  $c$ . In order to solve the game, we want to find a set of mixed strategies  $(p_r, p_c)$  that are, in some sense, in equilibrium. It may be that some of the equilibrium mixed strategies assign probability 1 to some choices, in which case they are interpreted as pure strategies.

The natural starting point in a search for a solution concept is standard decision theory: we assume that each player has some probability beliefs about the strategies that the other player might choose and that each player chooses the strategy that maximizes his expected payoff.

Suppose for example that the payoff to Row is  $u_r(r, c)$  if Row plays  $r$  and Column plays  $c$ . We assume that Row has a subjective probability distribution over Column's choices which we denote by  $(\pi_c)$ ; see Chapter 4 for the fundamentals of the idea of subjective probability. Here  $\pi_c$  is supposed to indicate the probability, as envisioned by Row, that Column will make the choice  $c$ . Similarly, Column has some beliefs about Row's behavior that we can denote by  $(\pi_r)$ .

We allow each player to play a mixed strategy and denote Row's actual mixed strategy by  $(p_r)$  and Column's actual mixed strategy by  $(p_c)$ . Since Row makes his choice without knowing Column's choice, Row's probability that a particular outcome  $(r, c)$  will occur is  $p_r\pi_c$ . This is simply the (objective) probability that Row plays  $r$  times Row's (subjective) probability that Column plays  $c$ . Hence, Row's objective is to choose a probability distribution  $(p_r)$  that maximizes

$$\text{Row's expected payoff} = \sum_r \sum_c p_r \pi_c u_r(r, c).$$

Column, on the other hand, wishes to maximize

$$\text{Column's expected payoff} = \sum_c \sum_r p_c \pi_r u_c(r, c).$$

So far we have simply applied a standard decision-theoretic model to this game – each player wants to maximize his or her expected utility given his or her beliefs. Given my beliefs about what the other player might do, I choose the strategy that maximizes my expected utility.

### 5.3.2 Nash equilibrium

In the expected payoff formulas given at the end of the last subsection, Row's behavior — how likely he is to play each of his strategies represented by the probability distribution  $(p_r)$  and Column's beliefs about Row's behavior are represented by the (subjective) probability distribution  $(\pi_r)$ .

A natural consistency requirement is that each player's belief about the other player's choices coincides with the actual choices the other player intends to make. Expectations that are consistent with actual frequencies are sometimes called rational expectations. A Nash equilibrium is a certain kind of rational expectations equilibrium. More formally:

**Definition 5.3.1 (Nash Equilibrium in Mixed Strategies.)** A Nash equilibrium in

mixed strategies consists of probability beliefs  $(\pi_r, \pi_c)$  over strategies, and probability of choosing strategies  $(p_r, p_c)$ , such that:

1. the beliefs are correct:  $p_r = \pi_r$  and  $p_c = \pi_c$  for all  $r$  and  $c$ ; and,
2. each player is choosing  $(p_r)$  and  $(p_c)$  so as to maximize his expected utility given his beliefs.

In this definition a Nash equilibrium in mixed strategies is an equilibrium in actions and beliefs. In equilibrium each player correctly foresees how likely the other player is to make various choices, and the beliefs of the two players are mutually consistent.

A more conventional definition of a Nash equilibrium in mixed strategies is that it is a pair of mixed strategies  $(p_r, p_c)$  such that each agent's choice maximizes his expected utility, given the strategy of the other agent. This is equivalent to the definition we use, but it is misleading since the distinction between the beliefs of the agents and the actions of the agents is blurred. We've tried to be very careful in distinguishing these two concepts.

One particularly interesting special case of a Nash equilibrium in mixed strategies is a Nash equilibrium in pure strategies, which is simply a Nash equilibrium in mixed strategies in which the probability of playing a particular strategy is 1 for each player. That is:

**Definition 5.3.2 (Nash equilibrium in Pure Strategies.)** A Nash equilibrium in pure strategies is a pair  $(r^*, c^*)$  such that  $u_r(r^*, c^*) \geq u_r(r, c^*)$  for all Row strategies  $r$ , and  $u_c(r^*, c^*) \geq u_c(r^*, c)$  for all Column strategies  $c$ .

A Nash equilibrium is a minimal consistency requirement to put on a pair of strategies: if Row believes that Column will play  $c^*$ , then Row's best reply is  $r^*$  and similarly for Column. No player would find it in his or her interest to deviate unilaterally from a Nash equilibrium strategy.

If a set of strategies is not a Nash equilibrium then at least one player is not consistently thinking through the behavior of the other player. That is, one of the players must expect the other player not to act in his own self-interest – contradicting the original hypothesis of the analysis.

An equilibrium concept is often thought of as a “rest point” of some adjustment process. One interpretation of Nash equilibrium is that it is the adjustment process of “thinking through” the incentives of the other player. Row might think: “If I think that Column is going to play some strategy  $c_1$  then the best response for me is to play  $r_1$ . But if Column thinks that I will play  $r_1$ , then the best thing for him to do is to play some other strategy  $c_2$ . But if Column is going to play  $c_2$ , then my best response is to play  $r_2$ ...” and so on.

**Example 5.3.1 (Nash equilibrium of Battle of the Sexes)** The following game is known as the “Battle of the Sexes.” The story behind the game goes something like this. Rhonda Row and Calvin Column are discussing whether to take microeconomics or macroeconomics this semester. Rhonda gets utility 2 and Calvin gets utility 1 if they both take micro; the payoffs are reversed if they both take macro. If they take different courses, they both get utility 0.

Let us calculate all the Nash equilibria of this game. First, we look for the Nash equilibria in pure strategies. This simply involves a systematic examination of the best responses to various strategy choices. Suppose that Column thinks that Row will play Top. Column gets 1 from playing Left and 0 from playing Right, so Left is Column’s best response to Row playing Top. On the other hand, if Column plays Left, then it is easy to see that it is optimal for Row to play Top. This line of reasoning shows that (Top, Left) is a Nash equilibrium. A similar argument shows that (Bottom, Right) is a Nash equilibrium.

|        |                | Calvin       |               |
|--------|----------------|--------------|---------------|
|        |                | Left (micro) | Right (macro) |
| Rhonda | Top (micro)    | (2, 1)       | (0, 0)        |
|        | Bottom (macro) | (0, 0)       | (1, 2)        |

Table 5.4: Battle of the Sexes

We can also solve this game systematically by writing the maximization problem that each agent has to solve and examining the first-order conditions. Let  $(p_t, p_b)$  be the probabilities with which Row plays Top and Bottom, and define  $(p_l, p_r)$  in a similar

manner. Then Row's problem is

$$\max_{(p_t, p_b)} p_t[p_l 2 + p_r 0] + p_b[p_l 0 + p_r 1]$$

such that

$$\begin{aligned} p_t &\geq 0 \\ p_b &\geq 0. \end{aligned}$$

Let  $\lambda$ ,  $\mu_t$ , and  $\mu_b$  be the Kuhn-Tucker multipliers on the constraints, so that the Lagrangian takes the form:

$$\mathcal{L} = 2p_t p_l + p_b p_r - \lambda(p_t + p_b - 1) - \mu_t p_t - \mu_b p_b.$$

Differentiating with respect to  $p_t$  and  $p_b$ , we see that the Kuhn-Tucker conditions for Row are

$$\begin{aligned} 2p_l &= \lambda + \mu_t \\ p_r &= \lambda + \mu_b \end{aligned} \tag{5.1}$$

Since we already know the pure strategy solutions, we only consider the case where  $p_t > 0$  and  $p_b > 0$ . The complementary slackness conditions then imply that  $\mu_t = \mu_b = 0$ . Using the fact that  $p_l + p_r = 1$ , we easily see that Row will find it optimal to play a mixed strategy when  $p_l = 1/3$  and  $p_r = 2/3$ .

Following the same procedure for Column, we find that  $p_t = 2/3$  and  $p_b = 1/3$ . The expected payoff to each player from this mixed strategy can be easily computed by plugging these numbers into the objective function. In this case the expected payoff is  $2/3$  to each player. Note that each player would prefer either of the pure strategy equilibria to the mixed strategy since the payoffs are higher for each player.

**Remark 5.3.1** One disadvantage of the notion of a mixed strategy is that it is sometimes difficult to give a behavioral interpretation to the idea of a mixed strategy although a mixed strategies are the only sensible equilibrium for some games such as Matching Pennies. For example, a duopoly game – mixed strategies seem unrealistic.

### 5.3.3 Dominant strategies

Let  $r_1$  and  $r_2$  be two of Row's strategies. We say that  $r_1$  strictly dominates  $r_2$  for Row if the payoff from strategy  $r_1$  is strictly larger than the payoff for  $r_2$  no matter what choice Column makes. The strategy  $r_1$  weakly dominates  $r_2$  if the payoff from  $r_1$  is at least as large for all choices Column might make and strictly larger for some choice.

A **dominant strategy equilibrium** is a choice of strategies by each player such that each strategy (weakly) dominates every other strategy available to that player. One particularly interesting game that has a dominant strategy equilibrium is the Prisoner's Dilemma in which the dominant strategy equilibrium is (confess, confess). If I believe that the other agent will not confess, then it is to my advantage to confess; and if I believe that the other agent will confess, it is still to my advantage to confess.

Clearly, a dominant strategy equilibrium is a Nash equilibrium, but not all Nash equilibria are dominant strategy equilibria. A dominant strategy equilibrium, should one exist, is an especially compelling solution to the game, since there is a unique optimal choice for each player.

## 5.4 Repeated games

In many cases, it is not appropriate to expect that the outcome of a repeated game with the same players as simply being a repetition of the one-shot game. This is because the strategy space of the repeated game is much larger: each player can determine his or her choice at some point as a function of the entire history of the game up until that point. Since my opponent can modify his behavior based on my history of choices, I must take this influence into account when making my own choices.

Let us analyze this in the context of the simple Prisoner's Dilemma game described earlier. Here it is in the "long-run" interest of both players to try to get to the (Cooperate, Cooperate) solution. So it might be sensible for one player to try to "signal" to the other that he is willing to "be nice" and play cooperate on the first move of the game. It is in the short-run interest of the other player to Defect, of course, but is this really in his long-run interest? He might reason that if he defects, the other player may lose patience and simply play Defect himself from then on. Thus, the second player might lose in the

long run from playing the short-run optimal strategy. What lies behind this reasoning is the fact that a move that I make now may have repercussions in the future the other player's future choices may depend on my current choices.

Let us ask whether the strategy of (Cooperate, Cooperate) can be a Nash equilibrium of the repeated Prisoner's Dilemma. First we consider the case of where each player knows that the game will be repeated a fixed number of times. Consider the reasoning of the players just before the last round of play. Each reasons that, at this point, they are playing a one-shot game. Since there is no future left on the last move, the standard logic for Nash equilibrium applies and both parties Defect.

Now consider the move before the last. Here it seems that it might pay each of the players to cooperate in order to signal that they are "nice guys" who will cooperate again in the next and final move. But we've just seen that when the next move comes around, each player will want to play Defect. Hence there is no advantage to cooperating on the next to the last move — as long as both players believe that the other player will Defect on the final move, there is no advantage to try to influence future behavior by being nice on the penultimate move. The same logic of backwards induction works for two moves before the end, and so on. **In a repeated Prisoner's Dilemma with a finite number of repetitions, the Nash equilibrium is still to Defect in every round.**

The situation is quite different in a repeated game with an infinite number of repetitions. In this case, at each stage it is known that the game will be repeated at least one more time and therefore there will be some (potential) benefits to cooperation. Let's see how this works in the case of the Prisoner's Dilemma.

Consider a game that consists of an infinite number of repetitions of the Prisoner's Dilemma described earlier. The strategies in this repeated game are sequences of functions that indicate whether each player will Cooperate or Defect at a particular stage as a function of the history of the game up to that stage. The payoffs in the repeated game are the discounted sums of the payoffs at each stage; that is, if a player gets a payoff at time  $t$  of  $u_t$ , his payoff in the repeated game is taken to be  $\sum_{t=0}^{\infty} u_t / (1 + r)^t$ , where  $r$  is the discount rate.

Now we show that **as long as the discount rate is not too high there exists a Nash equilibrium pair of strategies such that each player finds it in his interest**



**to cooperate at each stage.** In fact, it is easy to exhibit an explicit example of such strategies. Consider the following strategy: “Cooperate on the current move unless the other player defected on the last move. If the other player defected on the last move, then Defect forever.” This is sometimes called a punishment strategy, for obvious reasons: if a player defects, he will be punished forever with a low payoff.

To show that a pair of punishment strategies constitutes a Nash equilibrium, we simply have to show that if one player plays the punishment strategy the other player can do no better than playing the punishment strategy. Suppose that the players have cooperated up until move  $T$  and consider what would happen if a player decided to Defect on this move. Using the numbers from the Prisoner’s Dilemma example, he would get an immediate payoff of 4, but he would also doom himself to an infinite stream of payments of 1. The discounted value of such a stream of payments is  $1/r$ , so his total expected payoff from Defecting is  $4 + 1/r$ .

On the other hand, his expected payoff from continuing to cooperate is  $3 + 3/r$ . Continuing to cooperate is preferred as long as  $3 + 3/r > 4 + 1/r$ , which reduces to requiring that  $r < 2$ . As long as this condition is satisfied, the punishment strategy forms a Nash equilibrium: if one party plays the punishment strategy, the other party will also want to play it, and neither party can gain by unilaterally deviating from this choice.

This construction is quite robust. Essentially the same argument works for any payoffs that exceed the payoffs from (Defect, Defect). A famous result known as the Folk Theorem asserts precisely this: in a repeated Prisoner’s Dilemma any payoff larger than the payoff received if both parties consistently defect can be supported as a Nash equilibrium.

**Example 5.4.1 ( Maintaining a Cartel)** Consider a simple repeated duopoly which yields profits  $(\pi_c, \pi_c)$  if both firms choose to play a Cournot game and  $(\pi_j, \pi_j)$  if both firms produce the level of output that maximizes their joint profits – that is, they act as a cartel. It is well-known that the levels of output that maximize joint profits are typically not Nash equilibria in a single-period game — each producer has an incentive to dump extra output if he believes that the other producer will keep his output constant. However, as long as the discount rate is not too high, the joint profit-maximizing solution will be a Nash equilibrium of the repeated game. The appropriate punishment strategy is for each firm to produce the cartel output unless the other firm deviates, in which case it

will produce the Cournot output forever. An argument similar to the Prisoner's Dilemma argument shows that this is a Nash equilibrium.

## 5.5 Refinements of Nash equilibrium

The Nash equilibrium concept seems like a reasonable definition of an equilibrium of a game. As with any equilibrium concept, there are two questions of immediate interest: 1) will a Nash equilibrium generally exist; and 2) will the Nash equilibrium be unique?

Existence, luckily, is not a problem. Nash (1950) showed that with a finite number of agents and a finite number of pure strategies, an equilibrium will always exist. It may, of course, be an equilibrium involving mixed strategies. We will shown in Chapter 7 that it always exists a pure strategy Nash equilibrium if the strategy space is a compact and convex set and payoffs functions are continuous and quasi-concave.

Uniqueness, however, is very unlikely to occur in general. We have already seen that there may be several Nash equilibria to a game. Game theorists have invested a substantial amount of effort into discovering further criteria that can be used to choose among Nash equilibria. These criteria are known as refinements of the concept of Nash equilibrium, and we will investigate a few of them below.

### 5.5.1 Elimination of dominated strategies

When there is no dominant strategy equilibrium, we have to resort to the idea of a Nash equilibrium. But typically there will be more than one Nash equilibrium. Our problem then is to try to eliminate some of the Nash equilibria as being “unreasonable.”

One sensible belief to have about players' behavior is that it would be unreasonable for them to play strategies that are dominated by other strategies. This suggests that when given a game, we should first eliminate all strategies that are dominated and then calculate the Nash equilibria of the remaining game. This procedure is called elimination of dominated strategies; it can sometimes result in a significant reduction in the number of Nash equilibria.

For example consider the game given in Table 5.5. Note that there are two pure strategy Nash equilibria, (Top, Left) and (Bottom, Right). However, the strategy Right

|     |        | Column |        |
|-----|--------|--------|--------|
|     |        | Left   | Right  |
| Row | Top    | (2, 2) | (0, 2) |
|     | Bottom | (2, 0) | (1, 1) |

Table 5.5: A Game With Dominated Strategies

weakly dominates the strategy Left for the Column player. If the Row agent assumes that Column will never play his dominated strategy, the only equilibrium for the game is (Bottom, Right).

Elimination of strictly dominated strategies is generally agreed to be an acceptable procedure to simplify the analysis of a game. Elimination of weakly dominated strategies is more problematic; there are examples in which eliminating weakly dominated strategies appears to change the strategic nature of the game in a significant way.

### 5.5.2 Sequential Games and Subgame Perfect Equilibrium

The games described so far in this chapter have all had a very simple dynamic structure: they were either one-shot games or a repeated sequence of one-shot games. They also had a very simple information structure: each player in the game knew the other player's payoffs and available strategies, but did not know in advance the other player's actual choice of strategies. Another way to say this is that up until now we have restricted our attention to games with simultaneous moves.

But many games of interest do not have this structure. In many situations at least some of the choices are made sequentially, and one player may know the other player's choice before he has to make his own choice. The analysis of such games is of considerable interest to economists since many economic games have this structure: a monopolist gets to observe consumer demand behavior before it produces output, or a duopolist may observe his opponent's capital investment before making its own output decisions, etc. The analysis of such games requires some new concepts.

Consider for example, the simple game depicted in Table 5.6. It is easy to verify that there are two pure strategy Nash equilibria in this game, (Top, Left) and (Bottom, Right). Implicit in this description of this game is the idea that both players make their choices

|     |        | Column |        |
|-----|--------|--------|--------|
|     |        | Left   | Right  |
| Row | Top    | (1, 9) | (1, 9) |
|     | Bottom | (0, 0) | (2, 1) |

Table 5.6: The Payoff Matrix of a Simultaneous-Move Game

simultaneously, without knowledge of the choice that the other player has made. But suppose that we consider the game in which Row must choose first, and Column gets to make his choice after observing Row's behavior.

In order to describe such a sequential game it is necessary to introduce a new tool, the game tree. This is simply a diagram that indicates the choices that each player can make at each point in time. The payoffs to each player are indicated at the "leaves" of the tree, as in Figure 5.1. This game tree is part of a description of the game in extensive form.

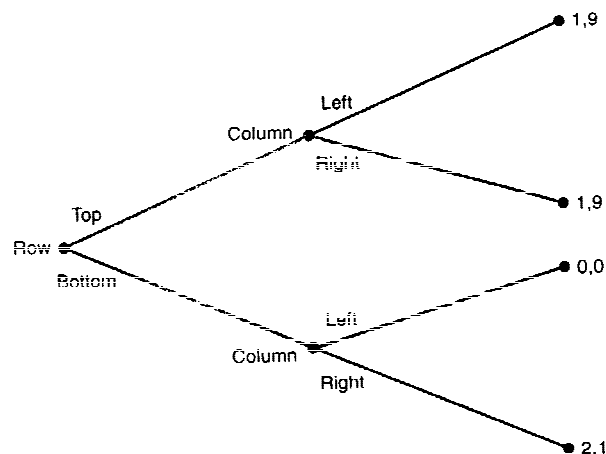


Figure 5.1: A game tree. This illustrates the payoffs to the previous game where Row gets to move first.

The nice thing about the tree diagram of the game is that it indicates the dynamic structure of the game — that some choices are made before others. A choice in the game corresponds to the choice of a branch of the tree. Once a choice has been made, the players are in a subgame consisting of the strategies and payoffs available to them from then on.

It is straightforward to calculate the Nash equilibria in each of the possible subgames,

particularly in this case since the example is so simple. If Row chooses top, he effectively chooses the very simple subgame in which Column has the only remaining move. Column is indifferent between his two moves, so that Row will definitely end up with a payoff of 1 if he chooses Top.

If Row chooses Bottom, it will be optimal for Column to choose Right, which gives a payoff of 2 to Row. Since 2 is larger than 1, Row is clearly better off choosing Bottom than Top. Hence the sensible equilibrium for this game is (Bottom, Right). This is, of course, one of the Nash equilibria in the simultaneous-move game. If Column announces that he will choose Right, then Row's optimal response is Bottom, and if Row announces that he will choose Bottom then Column's optimal response is Right.

But what happened to the other equilibrium, (Top, Left)? If Row believes that Column will choose Left, then his optimal choice is certainly to choose Top. But why should Row believe that Column will actually choose Left? Once Row chooses Bottom, the optimal choice in the resulting subgame is for Column to choose Right. A choice of Left at this point is not an equilibrium choice in the relevant subgame.

In this example, only one of the two Nash equilibria satisfies the condition that it is not only an overall equilibrium, but also an equilibrium in each of the subgames. A Nash equilibrium with this property is known as a **subgame perfect equilibrium**.

It is quite easy to calculate subgame-perfect equilibria, at least in the kind of games that we have been examining. One simply does a “backwards induction” starting at the last move of the game. The player who has the last move has a simple optimization problem, with no strategic ramifications, so this is an easy problem to solve. The player who makes the second to the last move can look ahead to see how the player with the last move will respond to his choices, and so on. The mode of analysis is similar to that of dynamic programming. Once the game has been understood through this backwards induction, the agents play it going forwards.

The extensive form of the game is also capable of modelling situations where some of the moves are sequential and some are simultaneous. The necessary concept is that of an **information set**. The information set of an agent is the set of all nodes of the tree that cannot be differentiated by the agent. For example, the simultaneous-move game depicted at the beginning of this section can be represented by the game tree in Figure

5.2. In this figure, the shaded area indicates that Column cannot differentiate which of these decisions Row made at the time when Column must make his own decision. Hence, it is just as if the choices are made simultaneously.

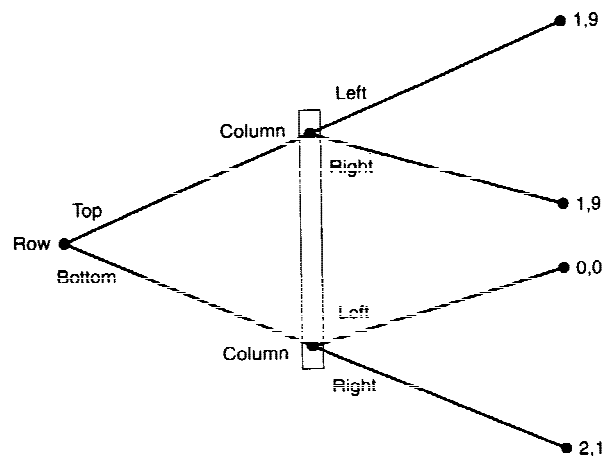


Figure 5.2: Information set. This is the extensive form to the original simultaneous-move game. The shaded information set indicates that column is not aware of which choice Row made when he makes his own decision.

Thus the extensive form of a game can be used to model everything in the strategic form plus information about the sequence of choices and information sets. In this sense the extensive form is a more powerful concept than the strategic form, since it contains more detailed information about the strategic interactions of the agents. It is the presence of this additional information that helps to eliminate some of the Nash equilibria as “unreasonable.”

**Example 5.5.1 (A Simple Bargaining Model)** Two players,  $A$  and  $B$ , have \$1 to divide between them. They agree to spend at most three days negotiating over the division. The first day,  $A$  will make an offer,  $B$  either accepts or comes back with a counteroffer the next day, and on the third day  $A$  gets to make one final offer. If they cannot reach an agreement in three days, both players get zero.

$A$  and  $B$  differ in their degree of impatience:  $A$  discounts payoffs in the future at a rate of  $\alpha$  per day, and  $B$  discounts payoffs at a rate of  $\beta$  per day. Finally, we assume that if a player is indifferent between two offers, he will accept the one that is most preferred by his opponent. This idea is that the opponent could offer some arbitrarily small amount

that would make the player strictly prefer one choice, and that this assumption allows us to approximate such an “arbitrarily small amount” by zero. It turns out that there is a unique subgame perfect equilibrium of this bargaining game.

As suggested above, we start our analysis at the end of the game, right before the last day. At this point  $A$  can make a take-it-or-leave-it offer to  $B$ . Clearly, the optimal thing for  $A$  to do at this point is to offer  $B$  the smallest possible amount that he would accept, which, by assumption, is zero. So if the game actually lasts three days,  $A$  would get 1 and  $B$  would get zero (i.e., an arbitrarily small amount).

Now go back to the previous move, when  $B$  gets to propose a division. At this point  $B$  should realize that  $A$  can guarantee himself 1 on the next move by simply rejecting  $B$ 's offer. A dollar next period is worth  $\alpha$  to  $A$  this period, so any offer less than  $\alpha$  would be sure to be rejected.  $B$  certainly prefers  $1 - \alpha$  now to zero next period, so he should rationally offer  $\alpha$  to  $A$ , which  $A$  will then accept. So if the game ends on the second move,  $A$  gets  $\alpha$  and  $B$  gets  $1 - \alpha$ .

Now move to the first day. At this point  $A$  gets to make the offer and he realizes that  $B$  can get  $1 - \alpha$  if he simply waits until the second day. Hence  $A$  must offer a payoff that has at least this present value to  $B$  in order to avoid delay. Thus he offers  $\beta(1 - \alpha)$  to  $B$ .  $B$  finds this (just) acceptable and the game ends. The final outcome is that the game ends on the first move with  $A$  receiving  $1 - \beta(1 - \alpha)$  and  $B$  receiving  $\beta(1 - \alpha)$ .

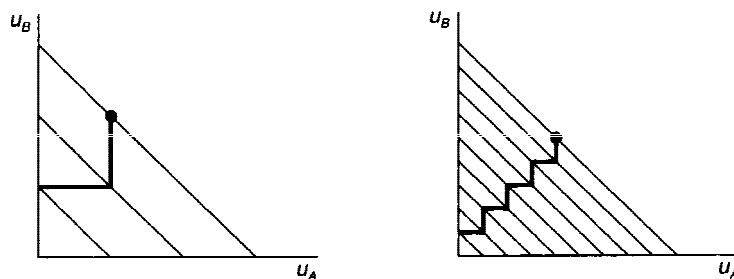


Figure 5.3: A bargaining game. The heavy line connects together the equilibrium outcomes in the subgames. The point on the outer-most line is the subgame-perfect equilibrium.

Figure 5.3 illustrates this process for the case where  $\alpha = \beta < 1$ . The outermost diagonal line shows the possible payoff patterns on the first day, namely all payoffs of the form  $x_A + x_B = 1$ . The next diagonal line moving towards the origin shows the

present value of the payoffs if the game ends in the second period:  $x_A + x_B = \alpha$ . The diagonal line closest to the origin shows the present value of the payoffs if the game ends in the third period; this equation for this line is  $x_A + x_B = \alpha^2$ . The right angled path depicts the minimum acceptable divisions each period, leading up to the final subgame perfect equilibrium. Figure 5.3B shows how the same process looks with more stages in the negotiation.

It is natural to let the horizon go to infinity and ask what happens in the infinite game. It turns out that the subgame perfect equilibrium division is payoff to  $A = \frac{1-\beta}{1-\alpha\beta}$  and payoff to  $B = \frac{\beta(1-\alpha)}{1-\alpha\beta}$ . Note that if  $\alpha = 1$  and  $\beta < 1$ , then player A receives the entire payoff, in accord with the principal expressed in the Gospels: “Let patience have her [subgame] perfect work.” (James 1:4).

### 5.5.3 Repeated games and subgame perfection

The idea of subgame perfection eliminates Nash equilibria that involve players threatening actions that are not credible - i.e., they are not in the interest of the players to carry out. For example, the Punishment Strategy described earlier is not a subgame perfect equilibrium. If one player actually deviates from the (Cooperate, Cooperate) path, then it is not necessarily in the interest of the other player to actually defect forever in response. It may seem reasonable to punish the other player for defection to some degree, but punishing forever seems extreme.

A somewhat less harsh strategy is known as Tit-for-Tat: start out cooperating on the first play and on subsequent plays do whatever your opponent did on the previous play. In this strategy, a player is punished for defection, but he is only punished once. In this sense Tit-for-Tat is a “forgiving” strategy.

Although the punishment strategy is not subgame perfect for the repeated Prisoner’s Dilemma, there are strategies that can support the cooperative solution that are subgame perfect. These strategies are not easy to describe, but they have the character of the West Point honor code: each player agrees to punish the other for defecting, and also punish the other for failing to punish another player for defecting, and so on. The fact that you will be punished if you don’t punish a defector is what makes it subgame perfect to carry out the punishments.



Unfortunately, the same sort of strategies can support many other outcomes in the repeated Prisoner's Dilemma. The Folk Theorem asserts that essentially all distributions of utility in a repeated one-shot game can be equilibria of the repeated game.

This excess supply of equilibria is troubling. In general, the larger the strategy space, the more equilibria there will be, since there will be more ways for players to "threaten" retaliation for defecting from a given set of strategies. In order to eliminate the "undesirable" equilibria, we need to find some criterion for eliminating strategies. A natural criterion is to eliminate strategies that are "too complex." Although some progress has been made in this direction, the idea of complexity is an elusive one, and it has been hard to come up with an entirely satisfactory definition.

## 5.6 Games with incomplete information

### 5.6.1 Bayes-Nash Equilibrium

Up until now we have been investigating games of complete information. In particular, each agent has been assumed to know the payoffs of the other player, and each player knows that the other agent knows this, etc. In many situations, this is not an appropriate assumption. If one agent doesn't know the payoffs of the other agent, then the Nash equilibrium doesn't make much sense. However, there is a way of looking at games of incomplete information due to Harsanyi (1967) that allows for a systematic analysis of their properties.

The key to the Harsanyi approach is to subsume all of the uncertainty that one agent may have about another into a variable known as the agent's type. For example, one agent may be uncertain about another agent's valuation of some good, about his or her risk aversion and so on. Each type of player is regarded as a different player and each agent has some prior probability distribution defined over the different types of agents.

A **Bayes-Nash equilibrium** of this game is then a set of strategies for each type of player that maximizes the expected value of each type of player, given the strategies pursued by the other players. This is essentially the same definition as in the definition of Nash equilibrium, except for the additional uncertainty involved about the type of the other player. Each player knows that the other player is chosen from a set of possible

types, but doesn't know exactly which one he is playing. Note in order to have a complete description of an equilibrium we must have a list of strategies for all types of players, not just the actual types in a particular situation, since each individual player doesn't know the actual types of the other players and has to consider all possibilities.

In a simultaneous-move game, this definition of equilibrium is adequate. In a sequential game it is reasonable to allow the players to update their beliefs about the types of the other players based on the actions they have observed. Normally, we assume that this updating is done in a manner consistent with Bayes' rule. Thus, if one player observes the other choosing some strategy  $s$ , the first player should revise his beliefs about what type the other player is by determining how likely it is that  $s$  would be chosen by the various types.

**Example 5.6.1 (A sealed-bid auction)** Consider a simple sealed-bid auction for an item in which there are two bidders. Each player makes an independent bid without knowing the other player's bid and the item will be awarded to the person with the highest bid. Each bidder knows his own valuation of the item being auctioned,  $v$ , but neither knows the other's valuation. However, each player believes that the other person's valuation of the item is uniformly distributed between 0 and 1. (And each person knows that each person believes this, etc.)

In this game, the type of the player is simply his valuation. Therefore, a Bayes-Nash equilibrium to this game will be a function,  $b(v)$ , that indicates the optimal bid,  $b$ , for a player of type  $v$ . Given the symmetric nature of the game, we look for an equilibrium where each player follows an identical strategy.

It is natural to guess that the function  $b(v)$  is strictly increasing; i.e., higher valuations lead to higher bids. Therefore, we can let  $V(b)$  be its inverse function so that  $V(b)$  gives us the valuation of someone who bids  $b$ . When one player bids some particular  $b$ , his probability of winning is the probability that the other player's bid is less than  $b$ . But this is simply the probability that the other player's valuation is less than  $V(b)$ . Since  $v$  is uniformly distributed between 0 and 1, the probability that the other player's valuation is less than  $V(b)$  is  $V(b)$ .

Hence, if a player bids  $b$  when his valuation is  $v$ , his expected payoff is

$$(v - b)V(b) + 0[1 - V(b)].$$

The first term is the expected consumer's surplus if he has the highest bid; the second term is the zero surplus he receives if he is outbid. The optimal bid must maximize this expression, so

$$(v - b)V'(b) - V(b) = 0.$$

For each value of  $v$ , this equation determines the optimal bid for the player as a function of  $v$ . Since  $V(b)$  is by hypothesis the function that describes the relationship between the optimal bid and the valuation, we must have

$$(V(b) - b)V'(b) \equiv V(b)$$

for all  $b$ .

The solution to this differential equation is

$$V(b) = b + \sqrt{b^2 + 2C},$$

where  $C$  is a constant of integration. In order to determine this constant of integration we note that when  $v = 0$  we must have  $b = 0$ , since the optimal bid when the valuation is zero must be 0. Substituting this into the solution to the differential equation gives us

$$0 = 0 + \sqrt{2C},$$

which implies  $C = 0$ . It follows that  $V(b) = 2b$ , or  $b = v/2$ , is a Bayes-Nash equilibrium for the simple auction. That is, it is a Bayes-Nash equilibrium for each player to bid half of his valuation.

The way that we arrived at the solution to this game is reasonably standard. Essentially, we guessed that the optimal bidding function was invertible and then derived the differential equation that it must satisfy. As it turned out, the resulting bid function had the desired property. One weakness of this approach is that it only exhibits one particular equilibrium to the Bayesian game — there could in principle be many others.

As it happens, in this particular game, the solution that we calculated is unique, but this need not happen in general. In particular, in games of incomplete information it may well pay for some players to try to hide their true type. For example, one type may try

to play the same strategy as some other type. In this situation the function relating type to strategy is not invertible and the analysis is much more complex.

### 5.6.2 Discussion of Bayesian-Nash equilibrium

The idea of Bayesian-Nash equilibrium is an ingenious one, but perhaps it is too ingenious. The problem is that the reasoning involved in computing Bayesian-Nash equilibria is often very involved. Although it is perhaps not unreasonable that purely rational players would play according to the Bayesian-Nash theory, there is considerable doubt about whether real players are able to make the necessary calculations.

In addition, there is a problem with the predictions of the model. The choice that each player makes depends crucially on his beliefs about the distribution of various types in the population. Different beliefs about the frequency of different types leads to different optimal behavior. Since we generally don't observe players beliefs about the prevalence of various types of players, we typically won't be able to check the predictions of the model. Ledyard (1986) has shown that essentially any pattern of play is a Bayesian-Nash equilibrium for some pattern of beliefs.

Nash equilibrium, in its original formulation, puts a consistency requirement on the beliefs of the agents – only those beliefs compatible with maximizing behavior were allowed. But as soon as we allow there to be many types of players with different utility functions, this idea loses much of its force. Nearly any pattern of behavior can be consistent with some pattern of beliefs.

## Reference

- Aumann, R. (1987). Game theory. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The New Palgrave*. (London: MacMillan Press).
- Binmore, K. (1991). *Fun and Games*. (San Francisco: Heath).
- Binmore, K. & Dasgupta, P. (1987). *The Economics of Bargaining*. (Oxford: Basil Blackwell).
- Fudenberg, D. & Tirole, J. (1991). *Game Theory*. (Cambridge: MIT Press).

- Kreps, D. (1990). *A Course in Microeconomic Theory*. (Princeton University Press).
- Harsanyi, J. (1967). Games of incomplete information played by Bayesian players. *Management Science*, 14, 159-182, 320-334, 486-502.
- Jehle, G. A., and P. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, 1998, Chapter 9.
- Ledyard, J. "The Scope of the Hypothesis of Bayesian Equilibrium," *Journal of Economic Theory*, 39 (1986), 59-82.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 8.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995, Chapters 8-9.
- Myerson, R. (1991). *Game Theory*. (Cambridge: Harvard University Press).
- Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, 54, 286-295.
- Rasmusen, E. (1989). *Games and Information*. (Oxford: Basil Blackwell).
- Tian, G. "On the Existence of Equilibria in Generalized Games," *International Journal of Game Theory*, 20 (1992), pp.247-254.
- Tian, G. "Existence of Equilibrium in Abstract Economies with Discontinuous Payoffs and Non-Compact Choice Spaces," *Journal of Mathematical Economics*, 21 (1992), pp. 379-388.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 15.

# Chapter 6

## Theory of the Market

### 6.1 Introduction

In previous chapters, we studied the behavior of individual consumers and firms, describing optimal behavior when markets prices were fixed and beyond the agent's control. Here, we explore the consequences of that behavior when consumers and firms come together in markets. We will consider equilibrium price and quantity determination in a single market or group of closed related markets by the actions of the individual agents for different markets structures. This equilibrium analysis is called a **partial equilibrium** analysis because it focus on a single market or group of closed related markets, implicitly assuming that changes in the markets under consideration do not change prices of other goods and upset the equilibrium that holds in those markets. We will treat all markets simultaneously in the general equilibrium theory.

We will concentrate on modelling the market behavior of the firm. How do firms determine the price at which they will sell their output or the prices at which they are willing to purchase inputs? We will see that in certain situations the “price-taking behavior” might be a reasonable approximation to optimal behavior, but in other situations we will have to explore models of the price-setting process. We will first consider the ideal (benchmark) case of perfect competition. We then turn to the study of settings in which some agents have market power. These settings include markets structures of pure monopoly, monopolistic competition, oligopoly, and monopsony.

## 6.2 The Role of Prices

The key insight of Adam Smith's *Wealth of Nations* is simple: if an exchange between two parties is voluntary, it will not take place unless both believe they will benefit from it. How is this also true for any number of parties and for production case? The price system is the mechanism that performs this task very well without central direction.

Prices perform three functions in organizing economic activities in a free market economy:

- (1) They transmit information about production and consumption. The price system transmits only the important information and only to the people who need to know. Thus, it transmits information in an efficient way.
- (2) They provide right incentives. One of the beauties of a free price system is that the prices that bring the information also provide an incentive to react on the information not only about the demand for output but also about the most efficient way to produce a product. They provide incentives to adopt those methods of production that are least costly and thereby use available resources for the most highly valued purposes.
- (3) They determine the distribution of income. They determine who gets how much of the product. In general, one cannot use prices to transmit information and provide an incentive to act on that information without using prices to affect the distribution of income. If what a person gets does not depend on the price he receives for the services of his resources, what incentive does he have to seek out information on prices or to act on the basis of that information?

## 6.3 Perfect Competition

Let us start to consider the case of pure competition in which there are a large number of independent sellers of some uniform product. In this situation, when each firm sets the price in which it sells its output, it will have to take into account not only the behavior of the consumers but also the behavior of the other producers.

### 6.3.1 Assumptions on Competitive Market

The competitive markets are based on the following assumptions:

- (1) Large number of buyers and sellers — price-taking behavior
- (2) Unrestricted mobility of resources among industries: no artificial barrier or impediment to entry or to exit from market.
- (3) Homogeneous product: All the firms in an industry produce an identical production in the consumers' eyes.
- (4) Passion of all relevant information (all relevant information are common knowledge): Firms and consumers have all the information necessary to make the correct economic decisions.

### 6.3.2 The Competitive Firm

A **competitive firm** is one that takes the market price of output as being given.

Let  $\bar{p}$  be the market price. Then the demand curve facing an ideal competitive firm takes the form

$$D(p) = \begin{cases} 0 & \text{if } p > \bar{p} \\ \text{any amount} & \text{if } p = \bar{p} \\ \infty & \text{if } p < \bar{p} \end{cases}$$

A competitive firm is free to set whatever price it wants and produce whatever quantity it is able to produce. However, if a firm is in a competitive market, it is clear that each firm that sells the product must sell it for the same price: for if any firm attempted to set its price at a level greater than the market price, it would immediately lose all of its customers. If any firm set its price at a level below the market price, all of the consumers would immediately come to it. Thus, each firm must take the market price as given, exogenous variable when it determines its supply decision.

### 6.3.3 The Competitive Firm's Short-Run Supply Function

Since the competitive firm must take the market price as given, its profit maximization problem is simple. The firm only needs to choose output level  $y$  so as to solve

$$\max_y py - c(y) \tag{6.1}$$



where  $y$  is the output produced by the firm,  $p$  is the price of the product, and  $c(y)$  is the cost function of production.

The first-order condition (in short, FOC) for interior solution gives:

$$p = c'(y) \equiv MC(y). \quad (6.2)$$

The first order condition becomes a sufficient condition if the second-order condition (in short, SOC) is satisfied

$$c''(y) > 0. \quad (6.3)$$

Taken together, these two conditions determine the supply function of a competitive firm: at any price  $p$ , the firm will supply an amount of output  $y(p)$  such that  $p = c'(y(p))$ . By  $p = c'(y(p))$ , we have

$$1 = c''(y(p))y'(p) \quad (6.4)$$

and thus

$$y'(p) > 0, \quad (6.5)$$

which means the law of supply holds.

Recall that  $p = c'(y^*)$  is the first-order condition characterizing the optimum only if  $y^* > 0$ , that is, only if  $y^*$  is an interior optimum. It could happen that at a low price a firm might very well decide to go out of business. For the short-run (in short, SR) case,

$$c(y) = c_v(y) + F \quad (6.6)$$

The firm should produce if

$$py(p) - c_v(y) - F \geq -F, \quad (6.7)$$

and thus we have

$$p \geq \frac{c_v(y(p))}{y(p)} \equiv AVC. \quad (6.8)$$

That is, the necessary condition for the firm to produce a positive amount of output is that the price is greater than or equal to the average variable cost.

Thus, the supply curve for the competitive firm is in general given by:  $p = c'(y)$  if  $p \geq \frac{c_v(y(p))}{y(p)}$  and  $y = 0$  if  $p \leq \frac{c_v(y(p))}{y(p)}$ . That is, the supply curve coincides with the upward sloping portion of the marginal cost curve as long as the price covers average variable cost, and the supply curve is zero if price is less than average variable cost.

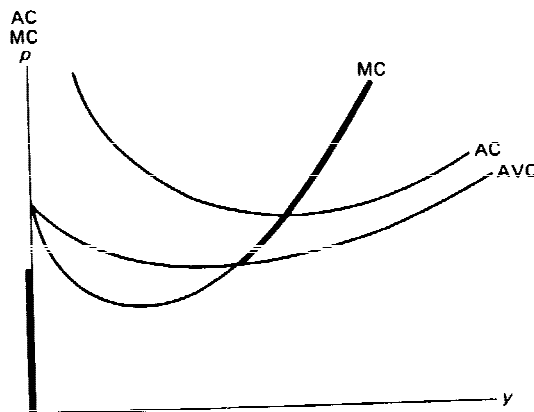


Figure 6.1: Firm's supply curve, and AC, AVC, and MC Curves.

Suppose that we have  $j$  firms in the market. (For the competitive model to make sense,  $j$  should be rather large.) The industry supply function is simply the sum of all individuals' supply functions so that it is given by

$$\hat{y}(p) = \sum_{j=1}^J y_j(p) \quad (6.9)$$

where  $y_i(p)$  is the supply function of firm  $j$  for  $j = 1, \dots, J$ . Since each firm chooses a level of output where price equals marginal cost, each firm that produces a positive amount of output must have the same marginal cost. The industry supply function measures the relationship between industry output and the common cost of producing this output.

The aggregate (industry) demand function measures the total output demanded at any price which is given by

$$\hat{x}(p) = \sum_{i=1}^n x_i(p) \quad (6.10)$$

where  $x_i(p)$  is the demand function of consumer  $i$  for  $i = 1, \dots, n$ .

### 6.3.4 Partial Market Equilibrium

How is the market price determined? It is determined by the requirement that the total amount of output that the firms wish to supply will be equal to the total amount of output that the consumers wish to demand. Formerly, we have

A *partial equilibrium price*  $p^*$  is a price where the aggregate quantity demanded equals

the aggregate quantity supplied. That is, it is the solution of the following equation:

$$\sum_{i=1}^n x_i(p) = \sum_{j=1}^J y_j(p) \quad (6.11)$$

Once this equilibrium price is determined, we can go back to look at the individual supply schedules of each firm and determine the firms level of output, its revenue, and its profits. In Figure 6.2, we have depicted cost curves for three firms. The first has positive profits, the second has zero profits, and the third has negative profits. Even though the third firm has negative profits, it may make sense for it to continue to produce as long as its revenues cover its variable costs.

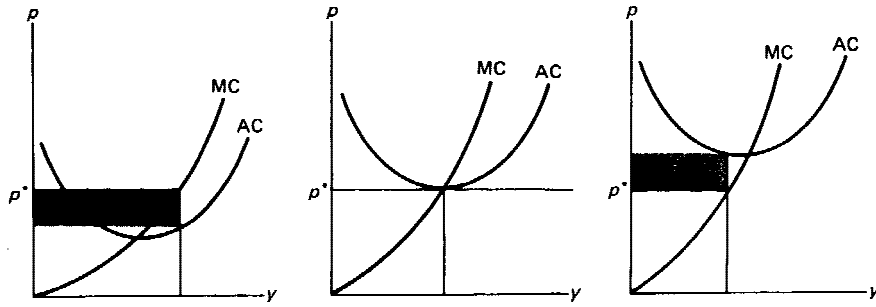


Figure 6.2: Positive, zero, and negative profits.

**Example 6.3.1**  $\hat{x}(p) = a - bp$  and  $c(y) = y^2 + 1$ . Since  $MC(y) = 2y$ , we have

$$y = \frac{p}{2} \quad (6.12)$$

and thus the industry supply function is

$$\hat{y}(p) = \frac{Jp}{2} \quad (6.13)$$

Setting  $a - bp = \frac{Jp}{2}$ , we have

$$p^* = \frac{a}{b + J/2}. \quad (6.14)$$

Now for general case of  $D(p)$  and  $S(p)$ , what happens about the equilibrium price if the number of firms increases? From

$$D(p(J)) = Jy(p(J))$$

we have

$$D'(p(J))p'(J) = y(p) + Jy'(p(J))p'(J)$$

and thus

$$p'(J) = \frac{y(p)}{X'(p) - Jy'(p)} < 0,$$

which means the equilibrium price decreases when the number of firms increases.

### 6.3.5 Competitive in the Long Run

The long-run behavior of a competitive industry is determined by two sorts of effects. The first effect is the free entry and exit phenomena so that the profits made by all firms are zero. If some firm is making negative profits, we would expect that it would eventually have to change its behavior. Conversely, if a firm is making positive profits we would expect that this would eventually encourage entry to the industry. If we have an industry characterized by free entry and exit, it is clear that in the long run all firms must make the same level of profits. As a result, every firm makes zero profit at the long-run competitive equilibrium.

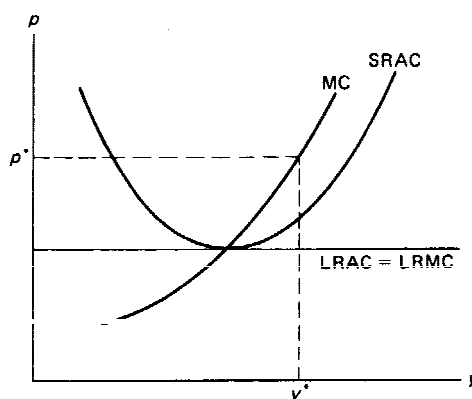


Figure 6.3: Long-run adjustment with constant costs.

The second influence on the long-run behavior of a competitive industry is that of technological adjustment. In the long run, firms will attempt to adjust their fixed factors so as to produce the equilibrium level of output in the cheapest way. Suppose for example we have a competitive firm with a long-run constant returns-to-scale technology that is operating at the position illustrated in Figure 6.3. Then in the long run it clearly pays

the firm to change its fixed factors so as to operate at a point of minimum average cost. But, if every firm tries to do this, the equilibrium price will certainly change.

In the model of entry or exit, the equilibrium number of firms is the largest number of firms that can break even so the price must be chosen to minimum price.

**Example 6.3.2**  $c(y) = y^2 + 1$ . The break-even level of output can be found by setting

$$AC(y) = MC(y)$$

so that  $y = 1$ , and  $p = MC(y) = 2$ .

Suppose the demand is linear:  $X(p) = a - bp$ . Then, the equilibrium price will be the smallest  $p^*$  that satisfies the conditions

$$p^* = \frac{a}{b + J/2} \geq 2.$$

As  $J$  increases, the equilibrium price must be closer and closer to 2.

## 6.4 Pure Monopoly

### 6.4.1 Profit Maximization Problem of Monopolist

At the opposite pole from pure competition we have the case of pure monopoly. Here instead of a large number of independent sellers of some uniform product, we have only one seller. A monopolistic firm must make two sorts of decisions: how much output it should produce, and at what price it should sell this output. Of course, it cannot make these decisions unilaterally. The amount of output that the firm is able to sell will depend on the price that it sets. We summarize this relationship between demand and price in a market demand function for output,  $y(p)$ . The market demand function tells how much output consumers will demand as a function of the price that the monopolist charges. It is often more convenient to consider the inverse demand function  $p(y)$ , which indicates the price that consumers are willing to pay for  $y$  amount of output. We have provided the conditions under which the inverse demand function exists in Chapter 2. The revenue that the firm receives will depend on the amount of output it chooses to supply. We write this revenue function as  $R(y) = p(y)y$ .

The cost function of the firm also depends on the amount of output produced. This relationship was extensively studied in Chapter 3. Here we have the factor prices as constant so that the conditional cost function can be written only as a function of the level of output of the firm.

The profit maximization problem of the firm can then be written as:

$$\max R(y) - c(y) = \max p(y)y - c(y)$$

The first-order conditions for profit maximization are that marginal revenue equals marginal cost, or

$$p(y^*) + p'(y^*)y^* = c'(y^*)$$

The intuition behind this condition is fairly clear. If the monopolist considers producing one extra unit of output he will increase his revenue by  $p(y^*)$  dollars in the first instance. But this increased level of output will force the price down by  $p'(y^*)$ , and he will lose this much revenue on unit of output sold. The sum of these two effects gives the marginal revenue. If the marginal revenue exceeds the marginal cost of production the monopolist will expand output. The expansion stops when the marginal revenue and the marginal cost balance out.

The first-order conditions for profit maximization can be expressed in a slightly different manner through the use of the price elasticity of demand. The price elasticity of demand is given by:

$$\epsilon(y) = \frac{p}{y(p)} \frac{dy(p)}{dp}$$

Note that this is always a negative number since  $dy(p)/dp$  is negative.

Simple algebra shows that the marginal revenue equals marginal cost condition can be written as:

$$p(y^*) \left[ 1 + \frac{y^*}{p(y^*)} \frac{dp(y^*)}{dy} \right] = p(y^*) \left[ 1 + \frac{1}{\epsilon(y^*)} \right] = c'(y^*)$$

that is, the price charged by a monopolist is a markup over marginal cost, with the level of the markup being given as a function of the price elasticity of demand.

There is also a nice graphical illustration of the profit maximization condition. Suppose for simplicity that we have a linear inverse demand curve:  $p(y) = a - by$ . Then the revenue

function is  $R(y) = ay - by^2$ , and the marginal revenue function is just  $R'(y) = a - 2by$ . The marginal revenue curve has the same vertical intercept as the demand curve but is twice as steep. We have illustrated these two curves in Figure 6.4, along with the average cost and marginal cost curves of the firm in question.

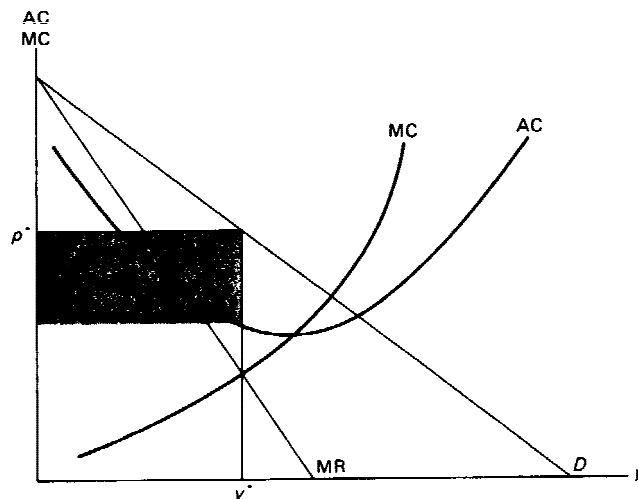


Figure 6.4: Determination of profit-maximizing monopolist's price and output.

The optimal level of output is located where the marginal revenue and the marginal cost curves intersect. This optimal level of output sells at a price  $p(y^*)$  so the monopolist gets an optimal revenue of  $p(y^*)y^*$ . The cost of producing  $y^*$  is just  $y^*$  times the average cost of production at that level of output. The difference between these two areas gives us a measure of the monopolist's profits.

### 6.4.2 Inefficiency of Monopoly

We say that a situation is Pareto efficient if there is no way to make one agent better off and the others are not worse off. Pareto efficiency will be a major theme in the discussion of welfare economics, but we can give a nice illustration of the concept here.

Let us consider the typical monopolistic configuration illustrated in Figure 6.5. It turns out a monopolist always operates in a Pareto inefficient manner. This means that there is some way to make the monopolist is better off and his customers are not worse off.

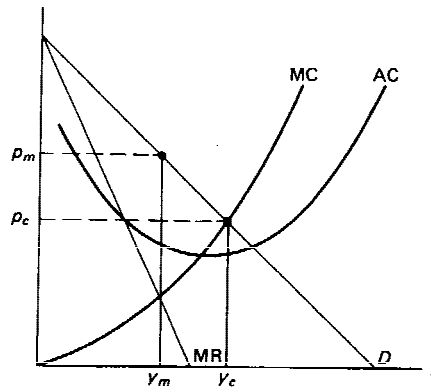


Figure 6.5: Monopoly results in Pareto inefficient outcome.

To see this let us think of the monopolist in Figure 6.5 after he has sold  $y_m$  of output at the price  $p_m$ , and received his monopolist profit. Suppose that the monopolist were to produce a small unit of output  $\Delta y$  more and offer to the public. How much would people be willing to pay for this extra unit? Clearly they would be willing to pay a price  $p(y_m + \Delta y)$  dollars. How much would it cost to produce this extra output? Clearly, just the marginal cost  $MC(y_m + \Delta y)$ . Under this rearrangement the consumers are at least not worse off since they are freely purchasing the extra unit of output, and the monopolist is better off since he can sell some extra units at a price that exceeds the cost of its production. Here we are allowing the monopolist to discriminate in his pricing: he first sells  $y_m$  and then sells more output at some other price.

How long can this process be continued? Once the competitive level of output is reached, no further improvements are possible. The competitive level of price and output is Pareto efficient for this industry. We will investigate the concept of Pareto efficiency in general equilibrium theory.

### 6.4.3 Monopoly in the Long Run

We have seen how the long-run and the short-run behavior of a competitive industry may differ because of changes in technology and entry. There are similar effects in a monopolized industry. The technological effect is the simplest: the monopolist will choose the level of his fixed factors so as to maximize his long-run profits. Thus, he will operate where marginal revenue equals long-run marginal cost, and that is all that needs to be



said.

The entry effect is a bit more subtle. Presumably, if the monopolist is earning positive profits, other firms would like to enter the industry. If the monopolist is to remain a monopolist, there must be some sort of barrier to entry so that a monopolist may make **positive profits** even in the long-run.

These barriers to entry may be of a legal sort, but often they are due to the fact that the monopolist owns some unique factor of production. For example, a firm might own a patent on a certain product, or might own a certain secret process for producing some item. If the monopoly power of the firm is due to a unique factor we have to be careful about we measure profits.

## 6.5 Monopolistic Competition

Recall that we assumed that the demand curve for the monopolist's product depended only on the price set by the monopolist. However, this is an extreme case. Most commodities have some substitutes and the prices of those substitutes will affect the demand for the original commodity. The monopolist sets his price assuming all the producers of other commodities will maintain their prices, but of course, this will not be true. The prices set by the other firms will respond — perhaps indirectly — to the price set by the monopolist in question. In this section we will consider what happens when several monopolists “compete” in setting their prices and output levels.

We imagine a group of  $n$  “monopolists” who sell similar, but not identical products. The price that consumers are willing to pay for the output of firm  $i$  depends on the level of output of firm  $i$  but also on the levels of output of the other firms: we write this inverse demand function as  $p_i(y_i, y)$  where  $y = (y_1 \dots y_n)$ .

Each firm is interested in maximizing profits: that is, each firm  $i$  wants to choose its level of output  $y_i$  so as to maximize:

$$p_i(y_i, y)y_i - c(y_i)$$

Unfortunately, the demand facing the  $i^{th}$  firm depends on what the other firms do. How is firm  $i$  supposed to forecast the other firms behavior?

We will adopt a very simple behavioral hypothesis: namely, that firm  $i$  assumes the other firms behavior will be constant. Thus, each firm  $i$  will choose its level of output  $y_i^*$  so as to satisfy:

$$p_i(y_i^*, y) + \frac{\partial p_i(y_i^*, y)}{\partial y_i} y_i^* - c'_i(y_i^*) \leq 0 \quad \text{with equality if } y_i^* > 0$$

For each combination of operating levels for the firms  $y_1, \dots, y_n$ , there will be some optimal operating level for firm  $i$ . We will denote this optimal choice of output by  $Y_i(y_1, \dots, y_n)$ . (Of course the output of firm  $i$  is not an argument of this function but it seems unnecessary to devise a new notation just to reflect that fact.)

In order for the market to be in equilibrium, each firm's forecast about the behavior of the other firms must be compatible with the other firms actually do. Thus if  $(y_1^*, \dots, y_n^*)$  is to be an equilibrium it must satisfy:

$$\begin{aligned} y_1^* &= Y_1(y_1^*, \dots, y_n^*) \\ &\vdots \\ y_n^* &= Y_n(y_1^*, \dots, y_n^*) \end{aligned}$$

that is,  $y_1^*$  must be the optimal choice for firm  $i$  if it assumes the other firms are going to produce  $y_2^*, \dots, y_n^*$ , and so on. Thus a monopolistic competition equilibrium  $(y_1^*, \dots, y_n^*)$  must satisfy:

$$p_i(y_i^*, y) + \frac{\partial p_i(y_i^*, y^*)}{\partial y_i} y_i^* - c'_i(y_i^*) \leq 0 \quad \text{with equality if } y_i^* > 0$$

and  $i = 1, \dots, n$ .

For each firm, its marginal revenue equals its marginal cost, given the actions of all the other firms. This is illustrated in Figure 6.6. Now, at the monopolistic competition equilibrium depicted in Figure 6.6, firm  $i$  is making positive profits. We would therefore expect others firms to enter the industry and share the market with the firm so the firm's profit will decrease because of close substitute goods. Thus, in the long run, firms would enter the industry until the profits of each firm were driven to zero. This means that firm  $i$  must charge a price  $p_i^*$  and produce an output  $y_i^*$  such that:

$$p_i^* y_i^* - c_i(y_i^*) \leq 0 \quad \text{with equality if } y_i^* > 0$$

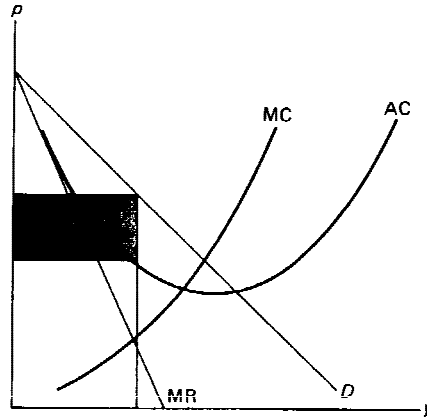


Figure 6.6: Short-run monopolistic competition equilibrium

or

$$p_i^* - \frac{c_i(y^*)}{y_i^*} \leq 0 \quad \text{with equality if } y_i^* > 0 \quad i = 1, 2, \dots, n.$$

Thus, the price must equal to average cost and on the demand curve facing the firm. As a result, as long as the demand curve facing each firm has some negative slope, each firm will produce at a point where average cost are greater than the minimum average costs. Thus, like a pure competitive firms, the profits made by each firm are zero and is very nearly the long run competitive equilibrium. On the other hand, like a pure monopolist, it still results in inefficient allocation as long as the demand curve facing the firm has a negative slope.

## 6.6 Oligopoly

Oligopoly is the study of market interactions with a small number of firms. Such an industry usually does not exhibit the characteristics of perfect competition, since individual firms' actions can in fact influence market price and the actions of other firms. The modern study of this subject is grounded almost entirely in the theory of games discussed in the last chapter. Many of the specifications of strategic market interactions have been clarified by the concepts of game theory. We now investigate oligopoly theory primarily from this perspective by introducing four models.

### 6.6.1 Cournot Oligopoly

A fundamental model for the analysis of oligopoly was the Cournot oligopoly model that was proposed by Cournot, an French economist, in 1838. A **Cournot equilibrium**, already mentioned in the last chapter, is a special set of production levels that have the property that no individual firm has an incentive to change its own production level if other firms do not change theirs.

To formalize this equilibrium concept, suppose there are  $J$  firms producing a single homogeneous product. If firm  $j$  produces output level  $q_j$ , the firm's cost is  $c_j(q_j)$ . There is a single market inverse demand function  $p(\hat{q})$ . The total supply is  $\hat{q} = \sum_{j=1}^J q_j$ . The profit to firm  $j$  is

$$p(\hat{q})q_j - c_j(q_j)$$

**Definition 6.6.1 (Cournot Equilibrium)** A set of output levels  $q_1, q_2, \dots, q_J$  constitutes a Cournot equilibrium if for each  $j = 1, 2, \dots, J$  the profit to firm  $j$  cannot be increased by changing  $q_j$  alone.

Accordingly, the Cournot model can be regarded as one shot game: the profit of firm  $j$  is its payoff, and the strategy space of firm  $j$  is the set of outputs, and thus a Cournot equilibrium is just a pure strategy Nash equilibrium. Then the first-order conditions for the interior optimum are:

$$p'(\hat{q})q_j + p(\hat{q}) - c'_j(q_j) = 0 \quad j = 1, 2, \dots, J.$$

The first-order condition for firm determines firm  $j$  optimal choice of output as a function of its beliefs about the sum of the other firms' outputs, denoted by  $\hat{q}_{-j}$ , i.e., the FOC condition can be written as

$$p'(q_j + \hat{q}_{-j})q_j + p(\hat{q}) - c'_j(q_j) = 0 \quad j = 1, 2, \dots, J.$$

The solution to the above equation, denoted by  $Q_j(\hat{q}_{-j})$ , is called the reaction function to the total outputs produced by the other firms.

Reaction functions give a direct characterization of a Cournot equilibrium. A set of output levels  $q_1, q_2, \dots, q_J$  constitutes a Cournot equilibrium if for each reaction function given  $q_j = Q_j(\hat{q}_{-j})$   $j = 1, 2, \dots, J$ .

An important special case is that of duopoly, an industry with just two firms. In this case, the reaction function of each firm is a function of just the other firm's output. Thus, the two reaction functions have the form  $Q_1(q_2)$  and  $Q_2(q_1)$ , which is shown in Figure 6.7. In the figure, if firm 1 selects a value  $q_1$  on the horizontal axis, firm 2 will react by selecting the point on the vertical axis that corresponds to the function  $Q_2(q_1)$ . Similarly, if firm 2 selects a value  $q_2$  on the vertical axis, firm 1 will be reacted by selecting the point on the horizontal axis that corresponds to the curve  $Q_1(q_2)$ . The equilibrium point corresponds to the point of intersection of the two reaction functions.

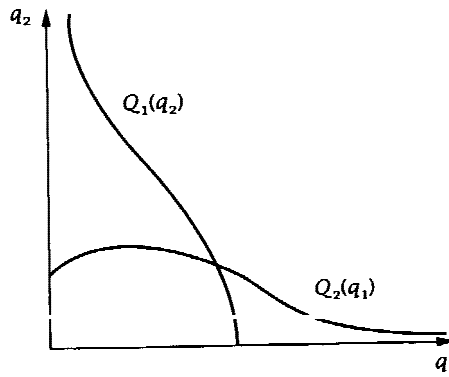


Figure 6.7: Reaction functions.

### 6.6.2 Stackelberg Model

There are alternative methods for characterizing the outcome of an oligopoly. One of the most popular of these is that of quantity leadership, also known as the Stackelberg model.

Consider the special case of a duopoly. In the Stackelberg formulation one firm, say firm 1, is considered to be the leader and the other, firm 2, is the follower. The leader may, for example, be the larger firm or may have better information. If there is a well-defined order for firms committing to an output decision, the leader commits first.

Given the committed production level  $q_1$  of firm 1, firm 2, the follower, will select  $q_2$  using the same reaction function as in the Cournot theory. That is, firm 2 finds  $q_2$  to maximize

$$\pi_2 = p(q_1 + q_2)q_2 - c_2(q_2),$$

where  $p(q_1 + q_2)$  is the industrywide inverse demand function. This yields the reaction function  $Q_2(q_1)$ .

Firm 1, the leader, accounts for the reaction of firm 2 when originally selecting  $q_1$ . In particular, firm 1 selects  $q_1$  to maximize

$$\pi_1 = p(q_1 + Q_2(q_1))q_1 - c_1(q_1),$$

That is, firm 1 substitutes  $Q_2(q_1)$  for  $q_2$  in the profit expression.

Note that a Stackelberg equilibrium does not yield a system of equations that must be solved simultaneously. Once the reaction function of firm 2 is found, firm 1's problem can be solved directly. Usually, the leader will do better in a Stackelberg equilibrium than in a Cournot equilibrium.

### 6.6.3 Bertrand Model

Another model of oligopoly of some interest is the so-called Bertrand model. The Cournot model and Stackelberg model take the firms' strategy spaces as being quantities, but it seems equally natural to consider what happens if price is chosen as the relevant strategic variables. Almost 50 years after Cournot, another French economist, Joseph Bertrand (1883), offered a different view of firm under imperfect competition and is known as the Bertrand model of oligopoly. Bertrand argued that it is much more natural to think of firms competing in their choice of price, rather than quantity. This small difference completely change the character of market equilibrium. This model is striking, and it contrasts starkly with what occurs in the Cournot model: With just two firms in a market, we obtain a perfectly competitive outcome in the Bertrand model!

In a simple Bertrand duopoly, two firms produce a homogeneous product, each has identical marginal costs  $c > 0$  and face a market demand curve of  $D(p)$  which is continuous, strictly decreasing at all price such that  $D(p) > 0$ . The strategy of each player is to announce the price at which he would be willing to supply an arbitrary amount of the good in question. In this case the payoff function takes a radically different form. It is plausible to suppose that the consumers will only purchase from the firm with the lowest price, and that they will split evenly between the two firms if they charge the same price.

This leads to a payoff to firm 1 of the form:

$$\pi_1(p_1, p_2) = \begin{cases} (p_1 - c)x(p_1) & \text{if } p_1 < p_2 \\ (p_1 - c)x(p_1)/2 & \text{if } p_1 = p_2 \\ 0 & \text{if } p_1 > p_2 \end{cases}.$$

Note that the Cournot game and the Bertrand game have a radically different structure. In the Cournot game, the payoff to each firm is a continuous function of its strategic choice; in the Bertrand game, the payoffs are discontinuous functions of the strategies. What is the Nash equilibrium? It may be somewhat surprising, but in the unique Nash equilibrium, both firms charge a price equal to marginal cost, and both earn zero profit. Formerly, we have

**Proposition 6.6.1** *There is a unique Nash equilibrium  $(p_1, p_2)$  in the Bertrand duopoly. In this equilibrium, both firms set their price equal to the marginal cost:  $p_1 = p_2 = c$  and earn zero profit.*

PROOF. First note that both firms setting their prices equal to  $c$  is indeed a Nash equilibrium. Neither firm can gain by raising its price because it will then make no sales (thereby still earning zero); and by lowering its price below  $c$  a firm increase its sales but incurs losses. What remains is to show that there can be no other Nash equilibrium. Because each firm  $i$  chooses  $p_i \geq c$ , it suffices to show that there are no equilibria in which  $p_i > c$  for some  $i$ . So let  $(p_1, p_2)$  be an equilibrium.

If  $p_1 > c$ , then because  $p_2$  maximizes firm 2's profits given firm 1's price choice, we must have  $p_2 \in (c, p_1]$ , because some such choice earns firm 2 strictly positive profits, whereas all other choices earn firm 2 zero profits. Moreover,  $p_2 \neq p_1$  because if firm 2 can earn positive profits by choosing  $p_2 = p_1$  and splitting the market, it can earn even higher profits by choosing  $p_2$  just slightly below  $p_1$  and supply the entire market at virtually the same price. Therefore,  $p_1 > c$  implies that  $p_2 > c$  and  $p_2 < p_1$ . But by stitching the roles of firms 1 and 2, an analogous argument establishes that  $p_2 > c$  implies that  $p_1 > c$  and  $p_1 < p_2$ . Consequently, if one firm's price is above marginal cost, both prices must be above marginal cost and each firm must be strictly undercutting the other, which is impossible.

### 6.6.4 Collusion

All of the models described up until now are examples of non-cooperative games. Each firm maximizes its own profits and makes its decisions independently of the other firms. What happens if they coordinate their actions? An industry structure where the firms collude to some degree in setting their prices and outputs is called a **cartel**.

A natural model is to consider what happens if the two firms choose simultaneously choose  $y_1$  and  $y_2$  to maximize industry profits:

$$\max_{y_1 y_2} p(y_1 + y_2)(y_1 + y_2) - c_1(y_1) - c_2(y_2)$$

The first-order conditions are

$$\begin{aligned} p'(y_1^* + y_2^*)(y_1^* + y_2^*) + p(y_1^* + y_2^*) &= c_1'(y_1^*) \\ p'(y_1^* + y_2^*)(y_1^* + y_2^*) + p(y_1^* + y_2^*) &= c_2'(y_2^*) \end{aligned}$$

It is easy to see from the above first-order conditions that profit maximization implies  $c_1'(y_1^*) = c_2'(y_2^*)$ .

The problem with the cartel solution is that it is not “stable” unless they are completely merged. There is always a temptation to cheat: to produce more than the agreed-upon output. Indeed, when the other firm will hold its production constant, we have by rearranging the first-order condition of firm 1:

$$\frac{\partial \pi_1(y^*, y_2^*)}{\partial y_1} = p'(y_1^* + y_2^*)y_1^* + p(y_1^* + y_2^*) - c_1'(y_1^*) = -p'(y_1^* + y_2^*)y_2^* > 0$$

by noting the fact that the demand curves slope downward.

The strategic situation is similar to the Prisoner’s Dilemma: if you think that other firm will produce its quota, it pays you to defect to produce more than your quota. And if you think that the other firm will not produce at its quota, then it will in general be profitable for you to produce more than your quota.

In order to make the cartel outcome viable, some punishment mechanism should be provided for the cheat on the cartel agreement, say using a repeated game as discussed in the previous chapter.



## 6.7 Monopsony

In our discussion up until now we have generally assumed that we have been dealing with the market for an output good. Output markets can be classified as “competitive” or “monopolistic” depending on whether firms take the market price as given or whether firms take the demand behavior of consumers as given.

There is similar classification for inputs markets. If firms take the factor prices as given, then we have competitive factor markets. If instead there is only one firm which demands some factor of production and it takes the supply behavior of its suppliers into account, then we say we have a monopsonistic factor market. The behavior of a monopsonist is analogous to that of a monopolist. Let us consider a simple example of a firm that is a competitor in its output market but is the sole purchaser of some input good. Let  $w(x)$  be the (inverse) supply curve of this factor of production. Then the profit maximization problem is:

$$\max p f(x) - w(x)x$$

The first-order condition is:

$$p f'(x^*) - w(x^*) - w'(x^*)x^* = 0$$

This just says that the marginal revenue product of an additional unit of the input good should be equal to its marginal cost. Another way to write the condition is:

$$p \frac{\partial f(x^*)}{\partial x} = w(x^*)[1 + 1/\epsilon]$$

where  $\epsilon$  is the price elasticity of supply. As  $\epsilon$  goes to infinity the behavior of a monopsonist approaches that of a pure competitor.

Recall that in Chapter 3, where we defined the cost function of a firm, we considered only the behavior of a firm with competitive factor markets. However, it is certainly possible to define a cost function for a monopsonistic firm. Suppose for example that  $x_i(\mathbf{w})$  is the supply curve for factor  $i$ . Then we can define:

$$\begin{aligned} c(y) = \min & \quad \sum w_i x_i(\mathbf{w}) \\ \text{s.t.} & \quad f(x(\mathbf{w})) = y \end{aligned}$$

This just gives us the minimum cost of producing  $y$ .

## Reference

- Jehle, G. A., and P. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, 1998, Chapter 6.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 3.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995, Chapter 12.
- Shapiro, C. (1989). Theories of oligopoly behavior. In R. Schmalensee & R. Willig (Eds.), *Handbook of Industrial Organization*, volume 1. (Amsterdam: North-Holland).
- Singh, N. & Vives, X. (1984). Price and quantity competition in a differentiated duopoly. *Rand Journal of Economics*, 15, 546-554.
- Sonnenschein, H. (1968). The dual of duopoly is complementary monopoly: or, two of Cournot's theories are one. *Journal of Political Economy*, 36, 316-318.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 14 and 16.

## Part III

# General Equilibrium Theory and Social Welfare

Part III is devoted to an examination of competitive market economies from a general equilibrium perspective at which all prices are variable and equilibrium requires that all markets clear.

The content of Part III is organized into four chapters. Chapters 7 and 8 constitute the heart of the general equilibrium theory. Chapter 7 presents the formal structure of the equilibrium model, introduces the notion of competitive equilibrium (or called Walrasian equilibrium). The emphasis is on positive properties of the competitive equilibrium. We will discuss the existence, uniqueness, and stability of a competitive equilibrium. We will also discuss a more general setting of equilibrium analysis, namely the abstract economy which includes the general equilibrium model as a special case. Chapter 8 discusses the normative properties of the competitive equilibrium by introducing the notion of Pareto efficiency. We examine the relationship between the competitive equilibrium and Pareto optimality. The core is concerned with the proof of the two fundamental theorems of welfare economics. Chapter 9 explores extensions of the basic analysis presented in Chapters 7 and 8. Chapter 9 covers a number of topics whose origins lie in normative theory. We will study the important core equivalence theorem that takes the idea of Walrasian equilibria as the limit of cooperative equilibria as markets grow large, fairness of allocation, and social choice theory. Chapter 10 applies the general equilibrium framework developed in Chapters 7 to 9 to economic situations involving the exchange and allocation of resources under conditions of uncertainty.

It is greatly acknowledged that some materials are drawn from the references in the end of each chapter.

# Chapter 7

## Positive Theory of Equilibrium: Existence, Uniqueness, and Stability

### 7.1 Introduction

One of the great achievements of economic theory in the last century is general equilibrium theory. General equilibrium theory is a branch of theoretical neoclassical economics. It seeks to explain the behavior of supply, demand and prices in a whole economy, by considering equilibrium in many markets simultaneously, unlike partial equilibrium theory which considers only one market at a time. Interaction between markets may result in a conclusion that is not obtained in a partial equilibrium framework. As such, it is a benchmark model to study market economy and so an abstraction from a real economy; it is proposed as being a useful model, both by considering equilibrium prices as long-term prices and by considering actual prices as deviations from equilibrium. The theory dates to the 1870s, particularly the work of French economist Léon Walras, and thus it is often called the Walrasian theory of market.

From a positive viewpoint, the general equilibrium theory is a theory of the determination of equilibrium prices and quantities in a system of perfectly competitive markets. The proof of existence of a general equilibrium is generally considered one of the most important and robust results of economic theory. The mathematical model of a competitive economy of L. Walras (1874-77) was conceived as an attempt to explain the state of equilibrium reached by a large number of small agents interaction through markets.

Walras himself perceived that the theory would be vacuous without a mathematical argument in support of the existence of at least one equilibrium state. However, for more than half a century the equality of the number of equations and of the number of unknowns of the Walrasian model remained the only, unconvincing remark made in favor of the existence of a competitive equilibrium. Study of the existence problem began in the early 1930s when Neisser (1932), Stackelberg (1933), Zeuthen (1933), and Schesinger (1935) identified some of its basic features and when Wald (1935, 1936a, 1936b) obtained its first solutions. After an interruption of some twenty years, the questions of existence of an economic equilibrium were taken up again by Arrow and Debreu (1954), McKenzie (1954, 1955), Gale (1955), and many others.

Although the first proof of existence of a competitive equilibrium was obtained by Wald, von Neumann (1928, 1937) turned out to be of greater importance of the subject. He proved a topological lemma which, in its reformulation by Kakutani (1941) as a fixed-point theorem for a correspondence, became the most powerful tool for the proof of existence of an economic equilibrium.

A *general equilibrium* is defined as a state where the aggregate demand does not exceed the aggregate supply for all markets. Thus, equilibrium prices are endogenously determined.

The general equilibrium approach has two central features:

- (1) It views the economy as a closed and inter-related system in which we must simultaneously determine the equilibrium values of all variables of interests (consider all markets together).
- (2) It aims at reducing the set of variables taken as exogenous to a small number of physical realities.

It is to predict the final consumption and production in the market mechanism.

The general equilibrium theory consists of five components:

1. Economic institutional environment (the fundamentals of the economy): economy that consists of consumption space, preferences, endowments of consumers, and production possibility sets of producers.

2. Economic institutional arrangement: It is the price mechanism in which a price is quoted for every commodity.
3. The behavior assumptions: price taking behavior for consumers and firms, utility maximization and profit maximization.
4. Predicting outcomes: equilibrium analysis — positive analysis such as existence, uniqueness, and stability.
5. Evaluating outcomes: normative analysis such as allocative efficiency of general equilibrium.

Questions to be answered in the general equilibrium theory.

- A. The existence and determination of a general equilibrium: What kinds of restrictions on economic environments (consumption sets, endowments, preferences, production sets) would guarantee the existence of a general equilibrium.
- B. Uniqueness of a general equilibrium: What kinds of restrictions on economic environments would guarantee a general equilibrium to be unique?
- C. Stability of a general equilibrium: What kinds of restrictions economic environments would guarantee us to find a general equilibrium by changing prices, especially raising the price if excess demand prevails and lowering it if excess supply prevails?
- D. Welfare properties of a general equilibrium: What kinds of restrictions on consumption sets, endowments, preferences, production sets would ensure a general equilibrium to be social optimal – Pareto efficient?

## **7.2 The Structure of General Equilibrium Model**

### **7.2.1 Economic Environments**

The fundamentals of the economy are economic institutional environments that are exogenously given and characterized by the following terms:

$n$ : the number of consumers

$N = \{1, \dots, n\}$ : the set of consumers

$J$ : the number of producers (firms)

$L$ : the number of (private) goods

$X_i \subset \mathbb{R}^L$ : the consumption space of consumer  $i = 1, \dots, n$ , which specifies the boundary of consumptions, collection of all individually feasible consumptions of consumer  $i$ . Some components of an element may be negative such as a labor supply;

$\succsim_i$ : preferences ordering (or  $u_i$  if a utility function exists) of consumer  $i = 1, \dots, n$ ;

*Remark:*  $\succsim_i$  is a preference ordering if it is reflexive ( $x_i \succsim_i x_i$ ), transitive ( $x_i \succsim_i x'_i$  and  $x'_i \succsim_i x''_i$  implies  $x_i \succsim_i x''_i$ ), and complete (for any pair  $x_i$  and  $x'_i$ , either  $x_i \succsim_i x'_i$  or  $x'_i \succsim_i x_i$ ). It can be shown that it can be represented by a continuous utility function if  $\succsim_i$  are continuous. The existence of general equilibrium can be obtained even when preferences are weakened to be non-complete or non-transitive.

$w_i \in X_i$ : initial endowment vector of consumer  $i$ .

$e_i = (X_i, \succsim_i, w_i)$ : the characteristic of consumer  $i$ .

$Y_j$ : production possibility set of firm  $j = 1, 2, \dots, J$ , which is the characteristic of producer  $j$ .

$y_j \in Y_j$ : a production plan,  $y_j^l > 0$  means  $y_j^l$  is output and  $y_j^l < 0$  means  $y_j^l$  is input. Most elements of  $y_j$  for a firm are zero.

Recall that there can be three types of returns about production scales: non-increasing returns to scale (i.e.,  $y_j \in Y_j$  implies that  $\alpha y_j \in Y_j$  for all  $\alpha \in [0, 1]$ ), non-decreasing (i.e.,  $y_j \in Y_j$  implies that  $\alpha y_j \in Y_j$  for all  $\alpha \geq 1$ ) returns to scale, and constant returns to scale i.e.,  $y_j \in Y_j$  implies that  $\alpha y_j \in Y_j$  for all  $\alpha \geq 0$ ). In other words, decreasing returns to scale implies any feasible input-output vector can be scaled down; increasing returns to scale implies any feasible input-output vector can be scaled up,



constant returns to scale implies the production set is the conjunction of increasing returns and decreasing returns. Geometrically, it is a cone.

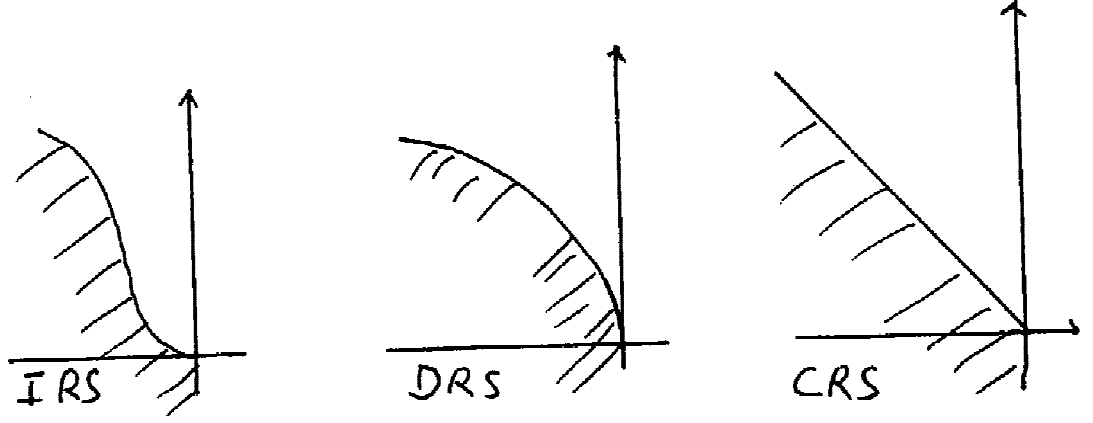


Figure 7.1: Various Returns to Scale: IRS, DRS, and CRS.

$e = (\{X_i, \succsim_i, w_i\}, \{Y_j\})$ : an economy, or called an economic environment.

$X = X_1 \times X_2 \times \dots \times X_n$ : consumption space.

$Y = Y_1 \times Y_2 \times \dots \times Y_J$ : production space.

## 7.2.2 Institutional Arrangement: Private Market Mechanism

$p = (p^1, p^2, \dots, p^L) \in \mathfrak{R}_+^L$ : a price vector;

$px_i$ : the expenditure of consumer  $i$  for  $i = 1, \dots, n$ ;

$py_j$ : the profit of firm  $j$  for  $j = 1, \dots, J$ ;

$pw_i$ : the value of endowments of consumer  $i$  for  $i = 1, \dots, n$ ;

$\theta_{ij} \in \mathfrak{R}_+$ : the profit share of consumer  $i$  from firm  $j$ , which specifies ownership (property rights) structures, so that  $\sum_{i=1}^n \theta_{ij} = 1$  for  $j = 1, 2, \dots, J$ ;

$\sum_{j=1}^J \theta_{ij} py_j$ : the total profit dividend received by consumer  $i$  from firms for  $i = 1, \dots, n$ .

For  $i = 1, 2, \dots, n$ , consumer  $i$ 's budget constraint is given by

$$px_i \leq pw_i + \sum_{j=1}^J \theta_{ij} py_j \quad (7.1)$$

and the budget set is given by

$$B_i(p) = \{x_i \in X_i : px_i \leq pw_i + \sum_{j=1}^J \theta_{ij} py_j\}. \quad (7.2)$$

A private ownership economy then is referred to

$$e = (e_1, e_2, \dots, e_n, \{Y_j\}_{j=1}^J, \{\theta_{ij}\}). \quad (7.3)$$

The set of all such private ownership economies are denoted by  $E$ .

### 7.2.3 Individual Behavior Assumptions:

- (1) Perfect competitive markets: Every player is a price-taker.
- (2) Utility maximization: Every consumer maximizes his preferences subject to  $B_i(p)$ . That is,

$$\max_{x_i} u_i(x_i) \quad (7.4)$$

s.t.

$$px_i \leq pw_i + \sum_{j=1}^J \theta_{ij} py_j \quad (7.5)$$

- (3) Profit maximization: Every firm maximizes its profit in  $Y_j$ . That is,

$$\max_{y_j \in Y_j} py_j \quad (7.6)$$

for  $j = 1, \dots, J$ .

### 7.2.4 Competitive Equilibrium

Before defining the notion of competitive equilibrium, we first give some notions on allocations which identify the set of possible outcomes in economy  $e$ . For notational convenience, “ $\hat{a}$ ” will be used throughout the notes to denote the sum of vectors  $a_i$ , i.e.,  $\hat{a} := \sum a_i$ .

**Allocation:**

An *allocation*  $(x, y)$  is a specification of consumption vector  $x = (x_1, \dots, x_n)$  and production vector  $y = (y_1, \dots, y_J)$ .

An allocation  $(x, y)$  is *individually feasible* if  $x_i \in X_i$  for all  $i \in N$ ,  $y_j \in Y_j$  for all  $j = 1, \dots, J$ .

An allocation is weakly balanced

$$\hat{x} \leq \hat{y} + \hat{w} \quad (7.7)$$

or specifically

$$\sum_{i=1}^n x_i \leq \sum_{j=1}^J y_j + \sum_{i=1}^n w_i \quad (7.8)$$

When inequality holds with equality, the allocation is called a balanced or attainable allocation.

An allocation  $(x, y)$  is feasible if it is both individually feasible and (weakly) balanced.

Thus, an economic allocation is feasible if the total amount of each good consumed does not exceed the total amount available from both the initial endowment and production. Denote by  $A = \{(x, y) \in X \times Y : \hat{x} \leq \hat{y} + \hat{w}\}$  the set of all feasible allocations.

**Aggregation:**

$\hat{x} = \sum_{i=1}^n x_i$ : aggregation of consumption;

$\hat{y} = \sum_{j=1}^J y_j$ : aggregation of production;

$\hat{w} = \sum_{i=1}^n w_i$ : aggregation of endowments;

Now we define the notion of competitive equilibrium.

**Definition 7.2.1 (Competitive Equilibrium or also called Walrasian Equilibrium)**

Given a private ownership economy,  $e = (e_1, \dots, e_n, \{Y_j\}, \{\theta_{ij}\})$ , an allocation  $(x, y) \in X \times Y$  and a price vector  $p \in \mathfrak{R}_+^L$  consist of a competitive equilibrium if the following conditions are satisfied

- (i) Utility maximization:  $x_i \succsim_i x'_i$  for all  $x'_i \in B_i(p)$  and  $x_i \in B_i(p)$  for  $i = 1, \dots, n$ .
- (ii) Profit maximization:  $py_j \geq py'_j$  for  $y'_j \in Y_j$ .
- (iii) Market Clear Condition:  $\hat{x} \leq \hat{w} + \hat{y}$ .

Denote

$x_i(p) = \{x_i \in B_i(p) : x_i \in B_i(p) \text{ and } x_i \succsim_i x'_i \text{ for all } x'_i \in B_i(p)\}$ : the demand correspondence of consumer  $i$  under utility maximization; it is called the demand function of consumer  $i$  if it is a single-valued function.

$y_j(p) = \{y_j \in Y_j : py_j \geq py'_j \text{ for all } y'_j \in Y_j\}$ : the supply correspondence of the firm  $j$ ; it is called the supply function of firm  $j$  if it is a single-valued function.

$\hat{x}(p) = \sum_{i=1}^n x_i(p)$  : the aggregate demand correspondence.

$\hat{y}(p) = \sum_{j=1}^J y_j(p)$  : the aggregate supply correspondence.

$\hat{z}(p) = \hat{x}(p) - \hat{w} - \hat{y}(p)$  : aggregate excess demand correspondence.

An equivalent definition of competitive equilibrium then is that a price vector  $p^* \in \mathbb{R}_+^L$  is a competitive equilibrium price if there exists  $\hat{z} \in \hat{z}(p^*)$  such that  $\hat{z} \leq 0$ .

If  $\hat{z}(p)$  is a single-valued,  $\hat{z}(p^*) \leq 0$  is a competitive equilibrium.

## 7.3 Some Examples of GE Models: Graphical Treatment

In most economies, there are three types of economic activities: production, consumption, and exchanges. Before formally stating the existence results on competitive equilibrium, we will first give two simple examples of general equilibrium models: exchange economies and a production economy with only one consumer and one firm. These examples introduce some of the questions, concepts, and common techniques that will occupy us for the rest of this part.

### 7.3.1 Pure Exchange Economies

A pure exchange economy is an economy in which there is no production. This is a special case of general economy. In this case, economic activities only consist of trading and consumption.

The aggregate excess demand correspondence becomes  $\hat{z}(p) = \hat{x}(p) - \hat{w}$  so that we can define the individual excess demand by  $z_i(p) = x_i(p) - w_i$  for this special case.

The simplest exchange economy with the possibility of mutual benefit exchange is the two-good and two-consumer exchange economy. As it turns out, this case is amenable to analysis by an extremely handy graphical device known as the Edgeworth Box.

- **Edgeworth Box:**

Consider an exchange economy with two goods ( $x_1, x_2$ ) and two persons. The total endowment is  $\hat{w} = w_1 + w_2$ . For example, if  $w_1 = (1, 2), w_2 = (3, 1)$ , then the total endowment is:  $\hat{w} = (4, 3)$ . Note that the point, denoted by  $w$  in the Edgeworth Box, can be used to represent the initial endowments of two persons.

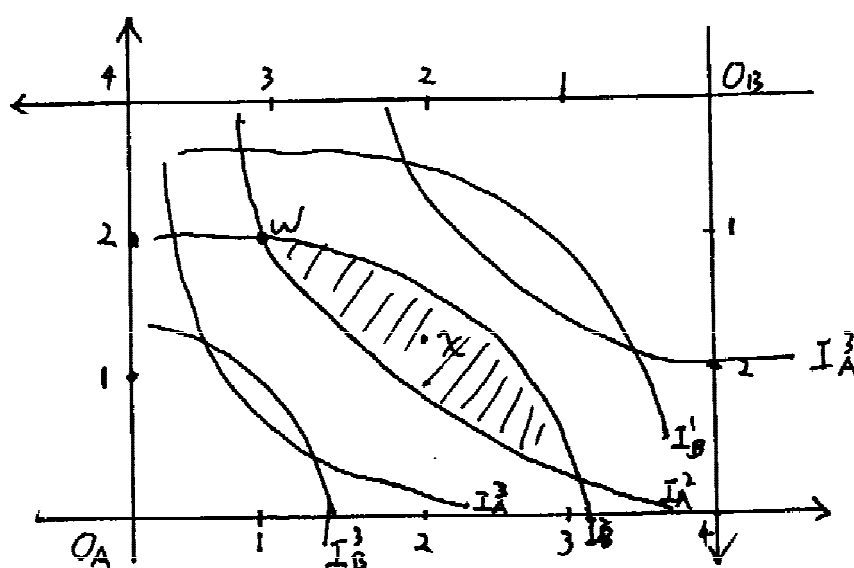


Figure 7.2: Edgeworth Box in which  $w_1 = (1, 2)$  and  $w_2 = (3, 1)$ .

Advantage of the Edgeworth Box is that it gives all the possible (balanced) trading points. That is,

$$x_1 + x_2 = w_1 + w_2 \quad (7.9)$$

for all points  $x = (x_1, x_2)$  in the box, where  $x_1 = (x_1^1, x_1^2)$  and  $x_2 = (x_2^1, x_2^2)$ . Thus, every point in the Edgeworth Box stands for an attainable allocation so that  $x_1 + x_2 = w_1 + w_2$ .

The shaded lens (portion) of the box in the above figure represents all the trading points that make both persons better off. Beyond the box, any point is not feasible.

Which point in the Edgeworth Box can be a competitive equilibrium?

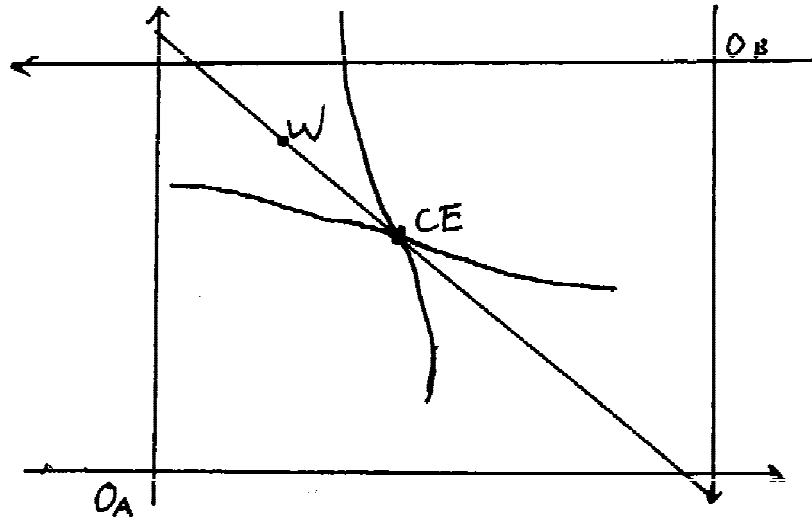


Figure 7.3: In the Edgeworth Box, the point CE is a competitive equilibrium.

In the box, one person's budget line is also the budget line of the other person. They share the same budget line in the box.

Figure 7.4 below shows the market adjustment process to a competitive equilibrium. Originally, at price  $p$ , both persons want more good 2. This implies that the price of good 1,  $p^1$ , is too high so that consumers do not consume the total amounts of the good so that there is a surplus for  $x_1$  and there is an excess demand for  $x_2$ , that is,  $x_1^1 + x_2^1 < w_1^1 + w_2^1$  and  $x_1^2 + x_2^2 > w_1^2 + w_2^2$ . Thus, the market will adjust itself by decreasing  $p^1$  to  $p^{1'}$ . As a result, the budget line will become flatter and flatter till it reaches the equilibrium where the aggregate demand equals the aggregate supply. In this interior equilibrium case, two indifference curves are tangent each other at a point that is on the budget line so that the marginal rates of substitutions for the two persons are the same that is equal to the price ratio.

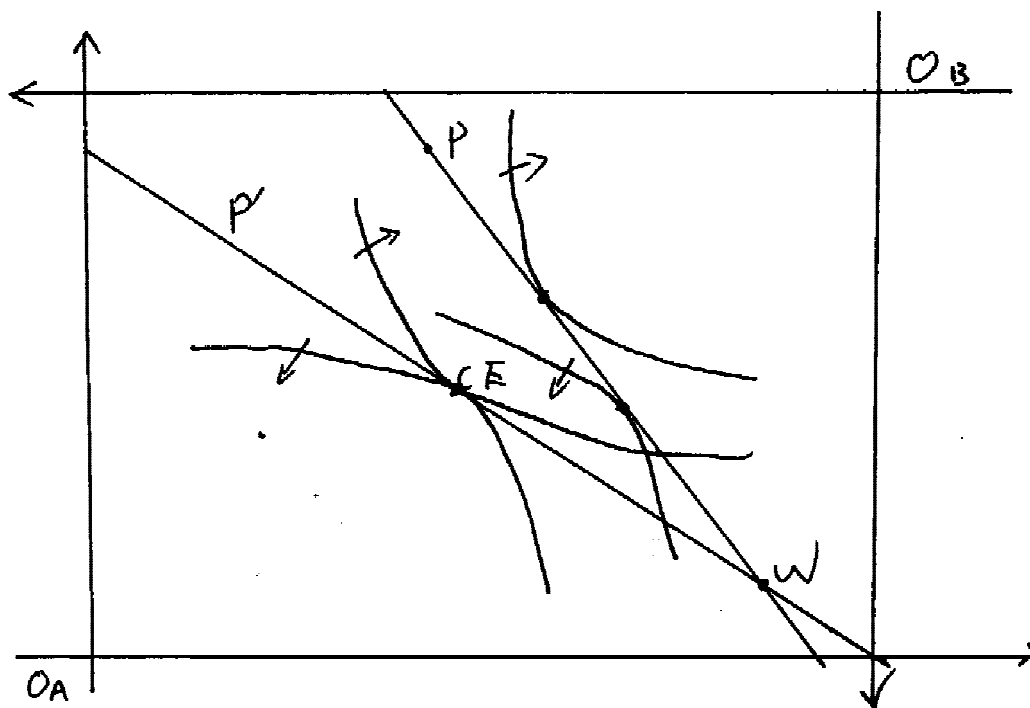


Figure 7.4: This figure shows the market adjustment process.

What happens when indifference curves are linear?

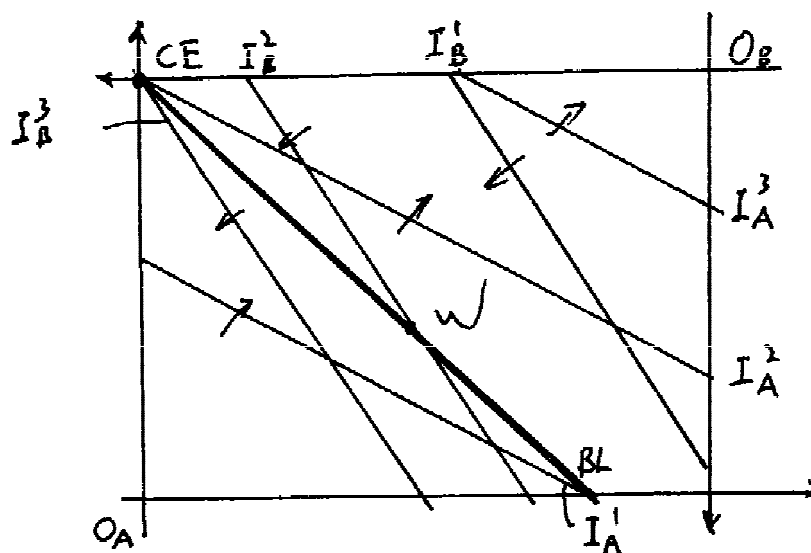


Figure 7.5: A competitive equilibrium may still exist even if two persons' indifference curves do not intersect.

In this case, there is no tangent point as long as the slopes of the indifference curves of the two persons are not the same. Even so, there still exists a competitive equilibrium although the marginal rates of substitutions for the two persons are not the same.

*Offer Curve*: the locus of the optimal consumptions for the two goods when price varies. Note that, it consists of tangent points of the indifference curves and budget lines when price varies.

The offer curves of two persons are given in the following figure:

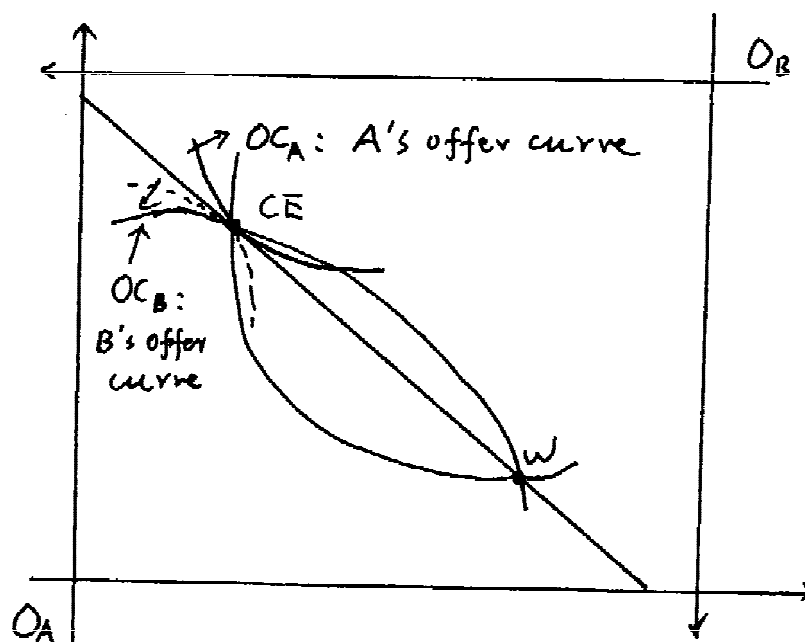


Figure 7.6: The CE is characterized by the intersection of two persons' offer curves.

The intersection of the two offer curves of the consumers can be used to check if there is a competitive equilibrium. By the above diagram, one can see that the one intersection of two consumers' offer curves is always given at the endowment  $w$ . If there is another intersection of two consumers' offer curves rather than the one at the endowment  $w$ , the intersection point must be the competitive equilibrium.

To enhance the understanding about the competitive equilibrium, you try to draw the competitive equilibrium in the following situations:

1. There are many equilibrium prices.



2. One person's preference is such that two goods are perfect substitutes (i.e., indifference curves are linear).
3. Preferences of one person are the Leontief-type (perfect complement)
4. One person's preferences are non-convex.
5. One person's preferences are "thick".
6. One person's preferences are convex, but has a satiation point.

Note that a preference relation  $\succsim_i$  is convex if  $x \succsim_i x'$  implies  $tx + (1-t)x' \succsim_i x'$  for all  $t \in (0, 1)$  and all  $x, x' \in X_i$ . A preference relation  $\succsim_i$  has a satiation point  $x$  if  $x \succsim_i x'$  all  $x' \in X_i$ .

*Cases in which there may be no Walrasian Equilibria:*

Case 1. Indifference curves (IC) are not convex. If the two offer curves are not intersected except for at the endowment points, then there may not exist a competitive equilibrium. This may be true when preferences are not convex.

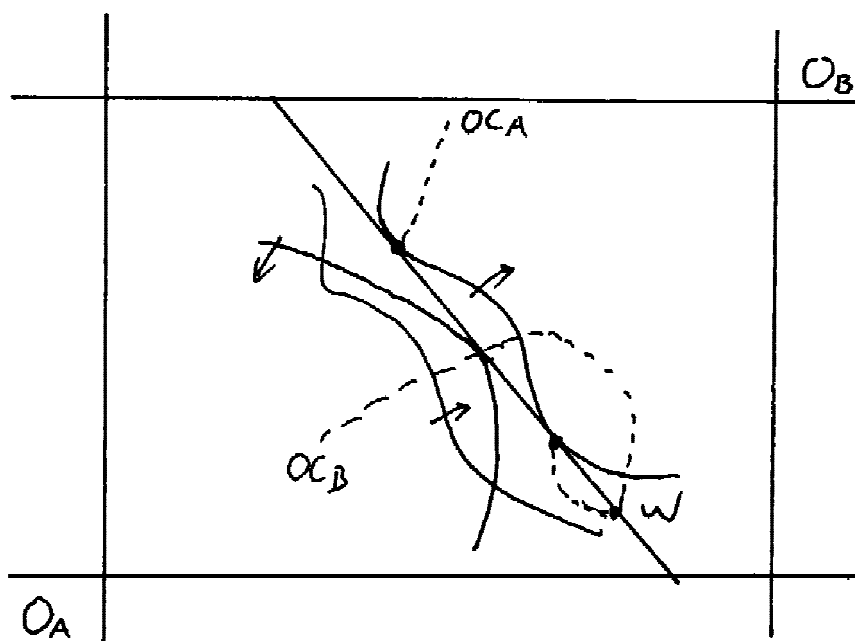


Figure 7.7: A CE may not exist if indifference curves are not convex. In that case, two persons' offer curves are not tangent.

Case 2. The initial endowment may not be an interior point of the consumption space. Consider an exchange economy in which one person's indifference curves put no

values on one commodity, say, when Person 2's indifference curves are vertical lines so that  $u_2(x_2^1, x_2^2) = x_2^1$ . Person 1's utility function is regular one, say, which is given by a quasi-linear utility function, say,  $u_1 = \sqrt{x_1^1} + x_1^2$ . The initial endowments are given by  $w_1 = (0, 1)$  and  $w_2 = (1, 0)$ . There is no competitive equilibrium. Why? There are two cases to be considered.

Case (a).  $p^1 = 0$  or  $p^2 = 0$ . Person 1 would demand infinite amount of the good whose price is zero. Thus, there is no competitive equilibrium in this case.

Case (b).  $p^1 > 0$  and  $p^2 > 0$ . Person 2's demand for good 1 is  $x_2^1 = 1$ , and person 1 would demand a positive quantity of good 1 because of the quasi-linear utility function. As such, the aggregate demand for good 1 is greater than one, that is,  $x_1^1(p) + x_2^1(p) > \hat{w}^1$  which violates the feasibility conditions. Thus, no CE exists.

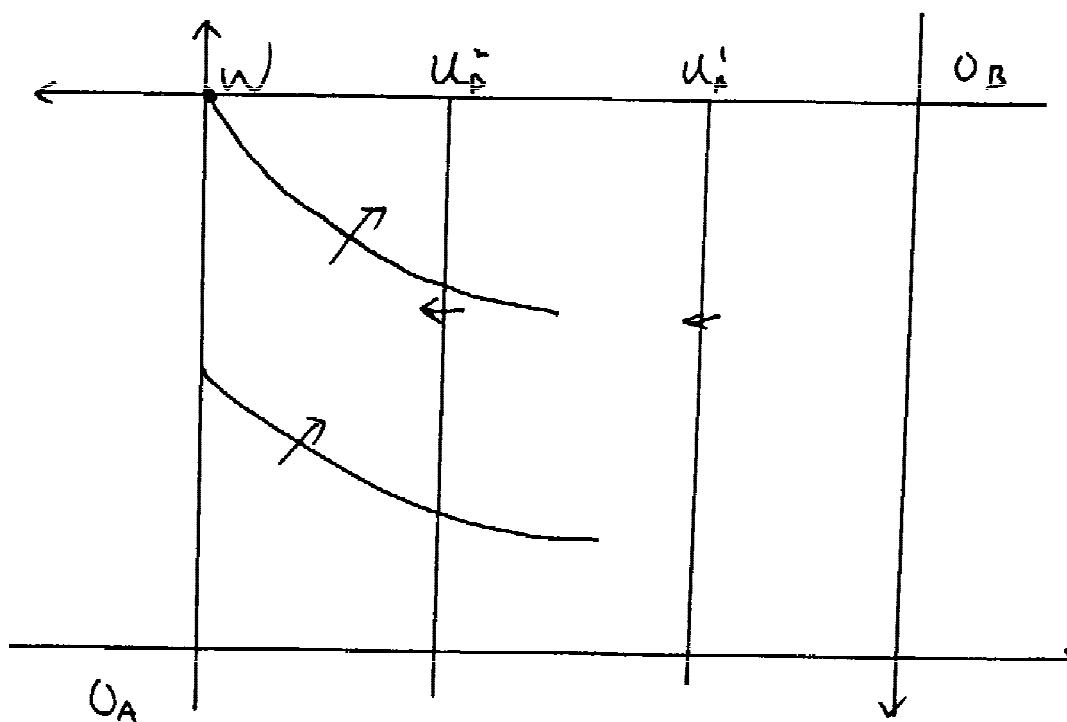


Figure 7.8: A CE may not exist if an endowment is on the boundary.

### 7.3.2 The One-Consumer and One Producer Economy

Now we introduce the possibility of production. To do so, in the simplest-possible setting in which there are only two price-taking economic agents.

Two agents: one producer so that  $J = 1$  and one consumer so that  $n = 1$ .

Two goods: labor (leisure) and the consumption good produced by the firm.

$w = (\bar{L}, 0)$ : the endowment.

$\bar{L}$ : the total units of leisure time.

$f(z) : \rightarrow R_+$ : the production function that is strictly increasing, concave, and differentiable, where  $z$  is labor input. To have an interior solution, we assume  $f$  satisfies the Inada condition  $f'(0) = +\infty$  and  $\lim_{z \rightarrow 0} f'(z)z = 0$ .

$(p, \omega)$ : the price vector of the consumption good and labor.

$\theta = 1$ : single person economy.

$u(x^1, x^2) : \rightarrow R_+^2$ : is the utility function which is strictly quasi-concave, increasing, and differentiable. To have an interior solution, we assume  $u$  satisfies the Inada condition  $\frac{\partial u}{\partial x_i}(0) = +\infty$  and  $\lim_{x_i \rightarrow 0} \frac{\partial u}{\partial x_i} x_i = 0$ .

The firm's problem is to choose the labor  $z$  so as to solve

$$\max_{z \geq 0} pf(z) - \omega z \quad (7.10)$$

FOC:

$$\begin{aligned} pf'(z) &= \omega \\ \Rightarrow f'(z) &= \omega/p \\ (MRTS_{z,q} &= \text{Price ratio}) \end{aligned}$$

which means the marginal rate of technique substitution of labor for the consumption good  $q$  equals the price ratio of the labor input and the consumption good output.

Let

$q(p, \omega)$  = the profit maximizing output for the consumption good.

$z(p, \omega)$  = the profit maximizing input for the labor.

$\pi(p, \omega)$  = the profit maximizing function.

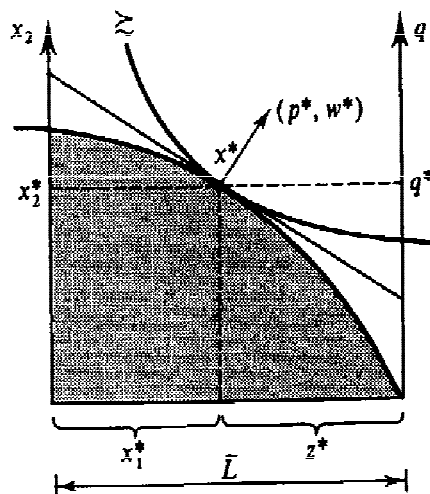
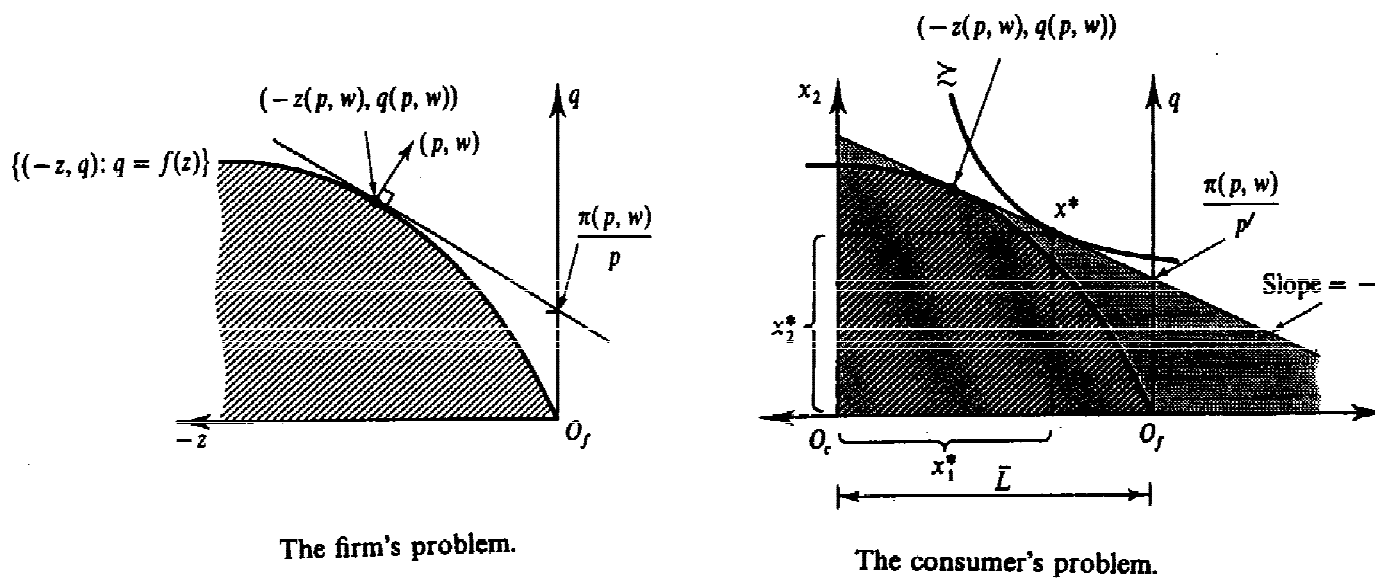


Figure 7.9: Figures for the producer's problem, the consumer problem and the CE.

The consumer's problem is to choose the leisure time and the consumption for the good so as to solve

$$\begin{aligned} & \max_{x^1, x^2} u(x^1, x^2) \\ \text{s.t.} \quad & px^2 \leq \omega(\bar{L} - x^1) + \pi(p, \omega) \end{aligned}$$

where  $x^1$  is the leisure and  $x^2$  is the consumption of good.

(FOC:)

$$\frac{\frac{\partial u}{\partial x^1}}{\frac{\partial u}{\partial x^2}} = \frac{\omega}{p} \quad (7.11)$$

which means the marginal rate of substitution of the leisure consumption for the consumption good  $q$  equals the price ratio of the leisure and the consumption good, i.e.,  $MRS_{x^1x^2} = \omega/p$ .

By (7.10) and (7.11)

$$MRS_{x^1x^2} = \frac{\omega}{p} = MRTS_{z,q} \quad (7.12)$$

A competitive equilibrium for this economy involves a price vector  $(p^*, \omega^*)$  at which

$$\begin{aligned} x^2(p^*, \omega^*) &= q(p^*, \omega^*); \\ x^1(p^*, \omega^*) + z(p^*, \omega^*) &= \bar{L} \end{aligned}$$

That is, the aggregate demand for the two goods equals the aggregate supply for the two goods. Figure 7.9 shows the problems of firm and consumer, and the competitive equilibrium, respectively.

## 7.4 The Existence of Competitive Equilibrium

The proof of the existence of a competitive equilibrium is generally considered one of the most important and robust results of economic theory. There are many different ways of establishing the existence of competitive equilibria, including the “excess demand approach” by showing that there is a price at which excess demand can be non-positive.

In this section we will examine the existence of competitive equilibrium for the three cases: (1) the single-valued aggregate excess demand function; (2) the aggregate excess demand correspondence; (3) a general class of private ownership production economies. The first two cases are based on excess demand instead of underlying preference orderings and consumption and production sets. There are many ways to prove the existence of general equilibrium. For instance, one can use the Brouwer fixed point theorem approach, KKM lemma approach, and abstract economy approach to show the existence of competitive equilibrium for these three cases.

### 7.4.1 The Existence of CE for Aggregate Excess Demand Functions

The simplest case for the existence of a competitive equilibrium is the one when the aggregate excess demand correspondence is a single-valued function. Note that, when preference orderings and production sets are both strictly convex, we obtain excess demand functions rather than correspondences.

A very important property of excess demand function  $\hat{z}(p)$  is Walras' law, which can take one of the following three forms:

(1) the strong Walras' law given by

$$p \cdot \hat{z}(p) = 0 \quad \text{for all } p \in \mathbb{R}_+^L;$$

(2) the weak Walras' law given by

$$p \cdot \hat{z}(p) \leq 0 \quad \text{for all } p \in \mathbb{R}_+^L;$$

(3) the interior Walras' law given by

$$p \cdot \hat{z}(p) = 0 \quad \text{for all } p \in \mathbb{R}_{++}^L.$$

Another important property of excess demand function is homogeneity of  $\hat{z}(p)$ : it is homogeneous of degree 0 in price  $\hat{z}(\lambda p) = \hat{z}(p)$  for any  $\lambda > 0$ . From this property, we can normalize prices.

Because of homogeneity, for example, we can normalize a price vector as follows:

$$(1) \quad p^l = p^l / p^1 \quad l = 1, 2, \dots, L$$

$$(2) \quad p^l = p^l / \sum_{l=1}^L p^l.$$

Thus, without loss of generality, we can restrict our attention to the unit simplex:

$$S^{L-1} = \{p \in \mathbb{R}_+^L : \sum_{l=1}^L p^l = 1\}. \quad (7.13)$$

Then, we have the following theorem on the existence of competitive equilibrium.

**Theorem 7.4.1 (The Existence Theorem I)** *For a private ownership economy  $e = (\{X_i, w_i, \succsim_i\}, \{Y_j\}, \{\theta_{ij};\})$ , if  $\hat{z}(p)$  is a homogeneous of degree zero and continuous function and satisfies the strong Walras' Law, then there exists a competitive equilibrium, that is, there is  $p^* \in \mathfrak{R}_+^L$  such that*

$$\hat{z}(p^*) \leq 0 \quad (7.14)$$

Proof: Define a continuous function  $g : S^{L-1} \rightarrow S^{L-1}$  by

$$g^l(p) = \frac{p^l + \max\{0, \hat{z}^l(p)\}}{1 + \sum_{k=1}^L \max\{0, \hat{z}^k(p)\}} \quad (7.15)$$

for  $l = 1, 2, \dots, L$ .

First note that  $g$  is a continuous function since  $\max\{f(x), h(x)\}$  is continuous when  $f(x)$  and  $h(x)$  are continuous.

By Brouwer's fixed point theorem, there exists a price vector  $p^*$  such that  $g(p^*) = p^*$ , i.e.,

$$p^{*l} = \frac{p^{*l} + \max\{0, \hat{z}^l(p^*)\}}{1 + \sum_{k=1}^L \max\{0, \hat{z}^k(p^*)\}} \quad l = 1, 2, \dots, L. \quad (7.16)$$

We want to show  $p^*$  is in fact a competitive equilibrium price vector.

Cross multiplying  $1 + \sum_{k=1}^L \max\{0, \hat{z}^k(p^*)\}$  on both sides of (7.16), we have

$$p^{*l} \sum_{k=1}^L \max\{0, \hat{z}^k(p^*)\} = \max\{0, \hat{z}^l(p^*)\}. \quad (7.17)$$

Then, multiplying the above equation by  $\hat{z}^l(p^*)$  and making summation, we have

$$\left[ \sum_{l=1}^L p^{*l} \hat{z}^l(p^*) \right] \left[ \sum_{l=1}^L \max\{0, \hat{z}^l(p^*)\} \right] = \sum_{l=1}^L \hat{z}^l(p^*) \max\{0, \hat{z}^l(p^*)\}. \quad (7.18)$$

Then, by the strong Walras' Law, we have

$$\sum_{l=1}^L \hat{z}^l(p^*) \max\{0, \hat{z}^l(p^*)\} = 0. \quad (7.19)$$

Therefore, each term of the summations is either zero or  $(\hat{z}^l(p^*))^2 > 0$ . Thus, to have the summation to be zero, we must have each term to be zero. That is,  $\hat{z}^l(p^*) \leq 0$  for  $l = 1, \dots, L$ . The proof is completed.

**Remark 7.4.1** Do not confuse competitive equilibrium of an aggregate excess demand function with the strong Walras' Law. Even though Walras' Law holds, we may not have

$\hat{z}(p) \leq 0$  for all  $p$ . Also, if  $\hat{z}(p^*) \leq 0$  for some  $p^*$ , i.e.,  $p^*$  is a competitive equilibrium price vector, the the strong Walras' Law may not hold unless some types of monotonicity are imposed such as local non-satiation.

**Fact 1: (free goods).** Under the strong Walras' Law, if  $p^*$  is a competitive equilibrium price vector and  $\hat{z}^l(p^*) < 0$ , then  $p^{*l} = 0$ .

Proof. Suppose not. Then  $p^{*l} > 0$ . Thus,  $p^{*l}\hat{z}^l(p^*) < 0$ , and so  $p^*\hat{z}(p^*) < 0$ , contradicting the strong Walras' Law.

*Desirable goods:* if  $p^l = 0$ ,  $\hat{z}^l(p^*) > 0$ .

**Fact 2:** (Equality of demand and supply). If all goods are desirable and  $p^*$  is a competitive equilibrium price vector, then  $\hat{z}(p^*) = 0$ .

Proof. Suppose not. We have  $\hat{z}^l(p^*) < 0$  for some  $l$ . Then, by Fact 1, we have  $p^{*l} = 0$ . Since good  $l$  is desirable, we must have  $\hat{z}^l(p^{*l}) > 0$ , a contradiction.

**Remark 7.4.2** By the strong Walras' Law, if  $p > 0$ , and if  $(L - 1)$  markets are in the equilibrium, the  $L$ -th market is also in the equilibrium. Thus, because of the strong Walras's Law, to verify that a price vector  $p > 0$  clears all markets, it suffices check that it clear all markets, but one.

The above result requires the aggregate excess demand function is continuous. By using the KKM lemma, we can prove the existence of competitive equilibrium by only assuming the aggregate excess demand function is lower semi-continuous.

**Theorem 7.4.2 (The Existence Theorem I')** *For a private ownership economy  $e = (\{X_i, w_i, \succsim_i\}, \{Y_j\}, \{\theta_{ij}\})$ , if the aggregate excess demand function  $\hat{z}(p)$  is a lower semi-continuous function and satisfies strong or weak Walras' Law, then there exists a competitive equilibrium, that is, there is  $p^* \in S^{L-1}$  such that*

$$\hat{z}(p^*) \leq 0. \quad (7.20)$$

Proof. Define a correspondence  $F : S^{L-1} \rightarrow 2^{S^{L-1}}$  by,

$$F(q) = \{p \in S^{L-1} : q\hat{z}(p) \leq 0\} \text{ for all } q \in S^{L-1}.$$

First note that  $F(q)$  is nonempty for each  $q \in S^{L-1}$  since  $q \in F(q)$  by strong or weak Walras' Law. Since  $p \geq 0$  and  $\hat{z}(\cdot)$  is lower semi-continuous, the function defined by



$\phi(q, p) \equiv q\hat{z}(p) = \sum_{l=1}^L q^l \hat{z}^l(p)$  is lower semi-continuous in  $p$ . Hence, the set  $F(q)$  is closed for all  $q \in S^{L-1}$ . We now prove  $F$  is FS-convex. Suppose, by way of contradiction, that there some  $q_1, \dots, q_m \in S^{L-1}$  and some convex combination  $q_\lambda = \sum_{t=1}^m \lambda_t q_t$  such that  $q_\lambda \notin \cup_{t=1}^m F(q_t)$ . Then,  $q_\lambda \notin F(q_t)$  and therefore  $q_t \hat{z}(q_\lambda) > 0$  for all  $t = 1, \dots, m$ . Thus,  $\sum_{t=1}^m \lambda_t q_t \hat{z}(q_\lambda) = q_\lambda \hat{z}(q_\lambda) > 0$  which contradicts the fact that  $\hat{z}$  satisfies strong or weak Walras' Law. So  $F$  must be FS-convex. Therefore, by KKM lemma, we have

$$\cap_{q \in S} F(q) \neq \emptyset.$$

Then there exists a  $p^* \in S^{L-1}$  such that  $p^* \in \cap_{q \in S^{L-1}} F(q)$ , i.e.,  $p^* \in F(q)$  for all  $q \in S^{L-1}$ . Thus,  $q\hat{z}(p^*) \leq 0$  for all  $q \in S^{L-1}$ . Now let  $q_1 = (1, 0, \dots, 0)$ ,  $q_2 = (0, 1, 0, \dots, 0)$ , and  $q_n = (0, \dots, 0, 1)$ . Then  $q_t \in S^{L-1}$  and thus  $q_t \hat{z}(p^*) = \hat{z}^t(p^*) \leq 0$  for all  $t = 1, \dots, L$ . Thus we have  $\hat{z}(p^*) \leq 0$ , which means  $p^*$  is a competitive equilibrium price vector. The proof is completed.

The above two existence theorems only give sufficient conditions. Recently, Tian (2013) provides a complete solution to the existence of competitive equilibrium in economies with general excess demand functions, in which commodities may be indivisible and excess demand functions may be discontinuous or do not have any structure except Walras' law.

Tian introduces a very weak condition, called *recursive transfer lower semi-continuity*, which is weaker than transfer lower semi-continuity and in turn weaker than lower semi-continuity. It is shown that the condition, together with Walras's law, guarantees the existence of price equilibrium in economies with excess demand functions. The condition is also necessary, and thus our results generalize all the existing results on the existence of price equilibrium in economies where excess demand is a function. For convenience of discussion, we introduce the following term.

We say that price system  $p$  *upsets* price system  $q$  if  $q$ 's excess demand is not affordable at price  $p$ , i.e.,  $p \cdot \hat{z}(q) > 0$ .

**Definition 7.4.1** An excess demand function  $\hat{z}(\cdot) : S^{L-1} \rightarrow \mathbb{R}^L$  is transfer lower semi-continuous if for all  $q, p \in S^{L-1}$ ,  $p \cdot \hat{z}(q) > 0$  implies that there exists some point  $p' \in X$  and some neighborhood  $\mathcal{N}(q)$  of  $q$  such that  $p' \cdot \hat{z}(q') > 0$  for all  $q' \in \mathcal{N}(q)$ .

**Remark 7.4.3** The transfer lower semi-continuity of  $\hat{z}(\cdot)$  means that, if the aggregate

excess demand  $\hat{z}(q)$  at price vector  $q$  is not affordable at price system  $p$ , then there exists a price system  $p'$  such that  $\hat{z}(q')$  is also not affordable at price system  $p'$ , provided  $q'$  is sufficiently close to  $q$ . Note that, since  $p \geq 0$ , this condition is satisfied if  $\hat{z}(\cdot)$  is lower semi-continuous by letting  $p' = p$ .

**Definition 7.4.2** (Recursive Upset Pricing) Let  $\hat{z}(\cdot) : S^{L-1} \rightarrow \mathbb{R}^L$  be an excess demand function. We say that a non-equilibrium price system  $p^0 \in S^{L-1}$  is *recursively upset by*  $p \in S^{L-1}$  if there exists a finite set of price systems  $\{p^1, p^2, \dots, p^m\}$  such that  $p^1 \cdot \hat{z}(p^0) > 0$ ,  $p^2 \cdot \hat{z}(p^1) > 0$ , ...,  $p \cdot \hat{z}(p^{m-1}) > 0$ .

In words, a non-equilibrium price system  $p^0$  is recursively upset by  $p$  means that there exist finite upsetting price systems  $p^1, p^2, \dots, p^m$  with  $p^m = p$  such that  $p^0$ 's excess demand is not affordable at  $p^1$ ,  $p^1$ 's excess demand is not affordable at  $p^2$ , and  $p^{m-1}$ 's excess demand is not affordable at  $p^m$ . When the strong form of Walras' law holds, this implies that  $p^0$  is upset by  $p^1$ ,  $p^1$  is upset by  $p^2$ , ...,  $p^{m-1}$  is upset by  $p$ .

For convenience, we say  $p^0$  is *directly upset by*  $p$  when  $m = 1$ , and *indirectly upset by*  $p$  when  $m > 1$ . Recursive upsetting says that nonequilibrium price system  $p^0$  can be directly or indirectly upset by a price system  $q$  through sequential upsetting price systems  $\{p^1, p^2, \dots, p^{m-1}\}$  in a recursive way that  $p^0$  is upset by  $p^1$ ,  $p^1$  is upset by  $p^2$ , ..., and  $p^{m-1}$  is upset by  $p$ .

**Definition 7.4.3** (Recursive Transfer Lower Semi-Continuity) An excess demand function  $\hat{z}(\cdot) : S^{L-1} \rightarrow \mathbb{R}^L$  is said to be *recursively transfer lower semi-continuous* on  $S^{L-1}$  if, whenever  $q \in S^{L-1}$  is not a competitive equilibrium price system, there exists some price system  $p^0 \in S^{L-1}$  (possibly  $p^0 = q$ ) and a neighborhood  $\mathcal{V}_q$  such that  $p \cdot \hat{z}(\mathcal{V}_q) > 0$  for any  $p$  that recursively upsets  $p^0$ , where  $p \cdot \hat{z}(\mathcal{V}_q) > 0$  means  $p \cdot \hat{z}(q') > 0$  for all  $q' \in \mathcal{V}_q$ .

In the definition of recursive transfer lower semi-continuity,  $q$  is transferred to  $q^0$  that could be any point in  $\Delta$ . Roughly speaking, recursive transfer lower semi-continuity of  $\hat{z}(\cdot)$  means that, whenever  $q$  is not an equilibrium price system, there exists another nonequilibrium price system  $p^0$  such that all excess demands in some neighborhood of  $q$  are not affordable at any price system  $p$  that recursively upsets  $p^0$ . This implies that, if an excess demand function  $\hat{z}(\cdot) : \Delta \rightarrow \mathbb{R}^L$  is not recursively transfer lower semi-continuous,

then there is some non-equilibrium price system  $q$  such that for every other price system  $p^0$  and every neighborhood of  $q$ , excess demand of some price system  $q'$  in the neighborhood becomes affordable at price system  $p$  that recursively upsets  $p^0$ .

**Remark 7.4.4** Recursive transfer lower semi-continuity is weaker than lower semi-continuity. Indeed, when  $\hat{z}(\cdot)$  is lower semi-continuous,  $p \cdot \hat{z}(\cdot)$  is also lower semi-continuous for any nonnegative vector  $p$ , and thus we have  $p \cdot \hat{z}(q') > 0$  for all  $q' \in \mathcal{V}_q$  and  $p \in S^{L-1}$ . Then, for any finite price vectors  $\{p^1, p^2, \dots, p^m\}$ , we of course have  $p^k \cdot \hat{z}(q') > 0$  for all  $q' \in \mathcal{V}_q$ ,  $k = 1, \dots, m$ , which means  $\hat{z}(\cdot)$  is transfer lower semi-continuous.

We then have the following theorem of competitive equilibrium in economies that have single-valued excess demand functions.

**Theorem 7.4.3 (The Existence Theorem I'')** *Suppose an excess demand function  $\hat{z}(\cdot) : S^{L-1} \rightarrow \mathbb{R}^L$  satisfies either the strong or weak Walras' law. Then there exists a competitive price equilibrium  $p^* \in S^{L-1}$  if and only if  $\hat{z}(\cdot)$  is recursively transfer lower semi-continuous on  $S^{L-1}$ .*

**Proof.** *Sufficiency* ( $\Leftarrow$ ). Suppose, by way of contradiction, that there is no price equilibrium. Then, by recursive transfer lower semi-continuity of  $\hat{z}(\cdot)$ , for each  $q \in S^{L-1}$ , there exists  $p^0$  and a neighborhood  $\mathcal{V}_q$  such that  $p \cdot \hat{z}(\mathcal{V}_q) > 0$  whenever  $p^0 \in S^{L-1}$  is recursively upset by  $p$ , i.e., for any sequence of recursive price systems  $\{p^1, \dots, p^{m-1}, p\}$  with  $p \cdot \hat{z}(p^{m-1}) > 0$ ,  $p^{m-1} \cdot \hat{z}(p^{m-2}) > 0$ ,  $\dots$ ,  $p^1 \cdot \hat{z}(p^0) > 0$  for  $m \geq 1$ , we have  $p \cdot \hat{z}(\mathcal{V}_q) > 0$ . Since there is no price equilibrium by the contrapositive hypothesis,  $p^0$  is not a price equilibrium and thus, by recursive transfer lower semi-continuity, such a sequence of recursive price systems  $\{p^1, \dots, p^{m-1}, p\}$  exists for some  $m \geq 1$ .

Since  $S^{L-1}$  is compact and  $S^{L-1} \subseteq \bigcup_{q \in S^{L-1}} \mathcal{V}_q$ , there is a finite set  $\{q^1, \dots, q^T\}$  such that  $S^{L-1} \subseteq \bigcup_{i=1}^T \mathcal{V}_{q^i}$ . For each of such  $q^i$ , the corresponding initial price system is denoted by  $p^{0i}$  so that  $p^i \cdot \hat{z}(\mathcal{V}_{q^i}) > 0$  whenever  $p^{0i}$  is recursively upset by  $p^i$ .

Since there is no price equilibrium, for each of such  $p^{0i}$ , there exists  $p^i$  such that  $p^i \cdot \hat{z}(p^{0i}) > 0$ , and then, by 1-recursive transfer lower semi-continuity, we have  $p^i \cdot \hat{z}(\mathcal{V}_{q^i}) > 0$ . Now consider the set of price systems  $\{p^1, \dots, p^T\}$ . Then,  $p^i \notin \mathcal{V}_{q^i}$ ; otherwise, by  $p^i \cdot \hat{z}(\mathcal{V}_{q^i}) > 0$ , we will have  $p^i \cdot \hat{z}(p^i) > 0$ , contradicting to Walras' law. So we must have  $p^1 \notin \mathcal{V}_{p^1}$ .

Without loss of generality, we suppose  $p^1 \in \mathcal{V}_{p^2}$ . Since  $p^2 \cdot \hat{z}(p^1) > 0$  by noting that  $p^1 \in \mathcal{V}_{q^2}$  and  $p^1 \cdot \hat{z}(p^{01}) > 0$ , then, by 2-recursive transfer lower semi-continuity, we have  $p^2 \cdot \hat{z}(\mathcal{V}_{q^1}) > 0$ . Also,  $q^2 \cdot \hat{z}(\mathcal{V}_{q^2}) > 0$ . Thus  $p^2 \cdot \hat{z}(\mathcal{V}_{q^1} \cup \mathcal{V}_{q^2}) > 0$ , and consequently  $p^2 \notin \mathcal{V}_{q^1} \cup \mathcal{V}_{q^2}$ .

Again, without loss of generality, we suppose  $p^2 \in \mathcal{V}_{q^3}$ . Since  $p^3 \cdot \hat{z}(p^2) > 0$  by noting that  $p^2 \in \mathcal{V}_{p^3}$ ,  $p^2 \cdot \hat{z}(p^1) > 0$ , and  $p^1 \cdot \hat{z}(p^{01}) > 0$ , by 3-recursive transfer lower semi-continuity, we have  $p^3 \cdot \hat{z}(\mathcal{V}_{q^1}) > 0$ . Also, since  $p^3 \cdot \hat{z}(p^2) > 0$  and  $p^2 \cdot \hat{z}(p^{02}) > 0$ , by 2-recursive transfer lower semi-continuity, we have  $p^3 \cdot \hat{z}(\mathcal{V}_{q^2}) > 0$ . Thus, we have  $p^3 \cdot \hat{z}(\mathcal{V}_{q^1} \cup \mathcal{V}_{q^2} \cup \mathcal{V}_{q^3}) > 0$ , and consequently  $p^3 \notin \mathcal{V}_{q^1} \cup \mathcal{V}_{q^2} \cup \mathcal{V}_{q^3}$ .

With this process going on, we can show that  $p^k \notin \mathcal{V}_{q^1} \cup \mathcal{V}_{q^2} \cup \dots \cup \mathcal{V}_{q^k}$ , i.e.,  $p^k$  is not in the union of  $\mathcal{V}_{q^1}, \mathcal{V}_{q^2}, \dots, \mathcal{V}_{q^k}$  for  $k = 1, 2, \dots, T$ . In particular, for  $k = T$ , we have  $p^L \notin \mathcal{V}_{q^1} \cup \mathcal{V}_{q^2} \dots \cup \mathcal{V}_{q^T}$  and so  $p^T \notin S^{L-1} \subseteq \mathcal{V}_{q^1} \cup \mathcal{V}_{q^2} \dots \cup \mathcal{V}_{q^T}$ , a contradiction.

Thus, there exists  $p^* \in S^{L-1}$  such that  $p \cdot \hat{z}(p^*) \leq 0$  for all  $p \in S^{L-1}$ . Letting  $p^1 = (1, 0, \dots, 0)$ ,  $p^2 = (0, 1, 0, \dots, 0)$ , and  $p^L = (0, 0, \dots, 0, 1)$ , we have  $\hat{z}^l(p^*) \leq 0$  for  $l = 1, \dots, L$  and thus  $p^*$  is a price equilibrium.

*Necessity* ( $\Rightarrow$ ). Suppose  $p^*$  is a competitive price equilibrium and  $p \cdot \hat{z}(q) > 0$  for  $q, p \in S^{L-1}$ . Let  $p^0 = p^*$  and  $N(q)$  be a neighborhood of  $q$ . Since  $p \cdot \hat{z}(p^*) \leq 0$  for all  $p \in S^{L-1}$ , it is impossible to find any sequence of finite price vectors  $\{p^1, p^2, \dots, p^m\}$  such that  $p^1 \cdot \hat{z}(p^0) > 0, p^2 \cdot \hat{z}(p^1) > 0, \dots, p^m \cdot \hat{z}(p^{m-1}) > 0$ . Hence, the recursive transfer lower semi-continuity holds trivially. ■

This theorem is useful to check the nonexistence of competitive equilibrium.

The method of proof employed to obtain Theorem 7.4.3 is new. While there are different ways of establishing the existence of competitive equilibrium, all the existing proofs essentially use the fixed-point-theorem related approaches. Moreover, a remarkable advantage of the above proof is that it is simple and elementary without using advanced math.

The above three existence theorems assume that the excess demand function is well defined for all prices in the closed unit simplex  $S^{L-1}$ , including zero prices. However, when preferences are strictly monotone, excess demand functions are not well defined on the boundary of  $S^{L-1}$  so that the above existence theorems cannot be applied. Then, we

need an existence theorem for strictly positive prices, which is given below.

**Theorem 7.4.4 (The Existence Theorem I''')** *For a private ownership economy  $e = (\{X_i, w_i, \succsim_i\}, \{Y_j\}, \{\theta_{ij}\})$ , suppose the aggregate excess demand function  $\hat{z}(p)$  is defined for all strictly positive price vectors  $p \in \mathfrak{R}_{++}^L$ , and satisfies the following conditions*

- (i)  $\hat{z}(\cdot)$  is continuous;
- (ii)  $\hat{z}(\cdot)$  is homogeneous of degree zero;
- (iii)  $p \cdot \hat{z}(p) = 0$  for all  $p \in \mathfrak{R}_{++}^L$  (interior Walras' Law);
- (iv) There is an  $s > 0$  such that  $\hat{z}^l(p) > -s$  for every commodity  $l$  and all  $p \in \mathfrak{R}_{++}^L$ ;
- (v) If  $p_k \rightarrow p$ , where  $p \neq 0$  and  $p^l = 0$  for some  $l$ , then

$$\max\{\hat{z}^1(p_k), \dots, \hat{z}^L(p_k)\} \rightarrow \infty.$$

Then there is  $p^* \in \mathfrak{R}_{++}^L$  such that

$$\hat{z}(p^*) = 0 \tag{7.21}$$

and thus  $p^*$  is a competitive equilibrium.

Proof. Because of homogeneity of degree zero we can restrict our search for an equilibrium in the unit simplex  $S$ . Denote its interior by  $\text{int}S$ . We want to construct a correspondence  $F$  from  $S$  to  $S$  such that any fixed point  $p^*$  of  $F$  is a competitive equilibrium, i.e.,  $p^* \in F(p^*)$  implies  $\hat{z}(p^*) = 0$ .

Define a correspondence  $F : S \rightarrow 2^S$  by,

$$F(p) = \begin{cases} \{q \in S : \hat{z}(p) \cdot q \geq \hat{z}(p) \cdot q' \text{ for all } q' \in S\} & \text{if } p \in \text{int}S \\ \{q \in S : p \cdot q = 0\} & \text{if } p \text{ is on the boundary} \end{cases}.$$

Note that, for  $p \in \text{int}S$ ,  $F(\cdot)$  means that, given the current "proposal"  $p \in \text{int}S$ , the "counterproposal" assigned by the correspondence  $F(\cdot)$  is any price vector  $q$  that maximizes the values of the excess demand vector among the permissible price vectors in  $S$ . Here  $F(\cdot)$  can be thought as a rule that adjusts current prices in a direction that eliminates any excess demand, the correspondence  $F(\cdot)$  assigns the highest prices to the commodities that are most in excess demand. In particular, we have

$$F(p) = \{q \in S : q^l = 0 \text{ if } \hat{z}^l(p) < \max\{\hat{z}^1(p), \dots, \hat{z}^L(p)\}\}.$$

Observe that if  $\hat{z}(p) \neq 0$  for  $p \in \text{int}S$ , then because of the interior Walras' law we have  $\hat{z}^l(p) < 0$  for some  $l$  and  $\hat{z}^{l'}(p) > 0$  for some  $l' \neq l$ . Thus, for such a  $p$ , any  $q \in F(p)$  has  $q^l = 0$  for some  $l$  (to maximize the values of excess demand vectors). Therefore, if  $\hat{z}(p) \neq 0$  then  $F(p) \subset \text{Boundary } S = S \setminus \text{Int}S$ . In contrast,  $\hat{z}(p) = 0$  then  $F(p) = S$ .

Now we want to show that the correspondence  $F$  is an upper hemi-continuous correspondence with nonempty, convex, and compact values. First, note that by the construction, the correspondence  $F$  is clearly compact and convex-valued.  $F(p)$  is also non-empty for all  $p \in S$ . Indeed, when  $p \in \text{int}S$ , any price vector  $q$  that maximizes the value of  $q' \cdot \hat{z}(p)$  is in  $F(p)$  and so it is not empty. When  $p$  is on the boundary of  $S$ ,  $p^l = 0$  for at least some good  $l$  and thus there exists a  $q \in S$  such that  $p \cdot q = 0$ , which implies  $F(p)$  is also non-empty.

Now we show the correspondence  $F$  is upper hemi-continuous, or equivalently it has closed graph, i.e., for any sequences  $p_t \rightarrow p$  and  $q_t \rightarrow q$  with  $q_t \in F(p_t)$  for all  $t$ , we have  $q \in F(p)$ . There are two cases to consider.

Case 1.  $p \in \text{int}S$ . Then  $p_k \in \text{int}S$  for  $k$  sufficiently large. From  $q_k \cdot \hat{z}(p_k) \geq q' \cdot \hat{z}(p_k)$  for all  $q' \in F(p_k)$  and the continuity of  $\hat{z}(\cdot)$ , we get  $q \cdot \hat{z}(p) \geq q' \cdot \hat{z}(p)$  for all  $q' \in F(p)$ , i.e.,  $q \in F(p)$ .

Case 2.  $p$  is a boundary point of  $S$ . Take any  $l$  with  $p^l > 0$ . We should argue that for  $k$  sufficiently large, we have  $q_k^l = 0$ , and therefore it must be that  $q^l = 0$ ; from this  $q \in F(p)$  follows. Because  $p^l > 0$ , there is an  $\epsilon > 0$  such that  $p_k^l > \epsilon$  for  $k$  sufficiently large. If, in addition,  $p_k$  is on the boundary of  $S$ , then  $q_k^l = 0$  by the definition of  $F(p_k)$ . If, instead,  $p_k \in \text{int}S$ , then by conditions (iv) and (v), for  $k$  sufficiently large, we must have

$$\hat{z}^l(p_k) < \max\{\hat{z}^1(p_k), \dots, \hat{z}^L(p_k)\}$$

and therefore that, again,  $q_k^l = 0$ . To prove the above inequality, note that by condition (v) the right-hand side of the above expression goes to infinity with  $k$  (because  $p$  is a boundary point of  $S$ , some prices go to zero as  $k \rightarrow \infty$ ). But the left-hand side is bounded above because if it is positive then

$$\hat{z}^l(p_k) \leq \frac{1}{\epsilon} p_k^l \hat{z}^l(p_k) = -\frac{1}{\epsilon} \sum_{l' \neq l} p_k^{l'} \hat{z}^{l'}(p_k) < \frac{s}{\epsilon} \sum_{l' \neq l} p_k^{l'} < \frac{s}{\epsilon},$$

where  $s$  is the bound in excess supply given by condition (iv). In summary, for  $p_k$  close e-

nough to the boundary of  $S$ , the maximal demand corresponds to some of the commodities whose price is close to zero. Therefore, we conclude that, for large  $k$ , any  $q_k \in F(p_k)$  will put nonzero weight only on commodities whose prices approach zero. But this guarantees  $p \cdot q = 0$  and so  $q \in F(p)$ . So  $F$  must be upper hemi-continuous.

Thus the correspondence  $F$  is an upper hemi-continuous correspondence with nonempty, convex, and compact values. Therefore, by Kakutani's fixed point theorem (see section 1.2.8), we conclude that there is  $p^* \in S$  with  $p^* \in F(p^*)$ .

Finally, we show that any fixed point  $p^*$  of  $F$  is a competitive equilibrium. Suppose that  $p^* \in F(p^*)$ . Then  $p^*$  cannot be a boundary point of  $S$  because  $p \cdot p > 0$  and  $p \cdot q = 0$  for all  $q \in F(p)$  cannot occur simultaneously, and thus  $p^* \in \text{int} S$ . If  $\hat{z}(p^*) \neq 0$ , then  $\hat{z}^l(p^*) < 0$  for  $l$  and  $\hat{z}^k(p^*) > 0$  for  $k$  by the interior Walras' Law. Thus, for such a  $p^*$ , any  $q^* \in F(p^*)$  must have  $q^{*l} = 0$  because  $q^*$  is the maximum of the function  $\hat{z}(p^*) \cdot q$ , which means  $F(p^*)$  is a subset of boundary points of  $S$ . Hence, if  $p^* \in F(p^*)$ , we must have  $\hat{z}(p^*) = 0$ . The proof is completed.

As shown by Tian (2009), Theorem 7.4.3 can also be extended to the case of any set, especially the positive price open set, of price systems for which excess demand is defined. To do so, we introduce the following version of recursive transfer lower semi-continuity.

**Definition 7.4.4** Let  $D$  be a subset of  $\text{int } S^{L-1}$ . An excess demand function  $\hat{z}(\cdot) : \text{int } S^{L-1} \rightarrow \mathbb{R}^L$  is said to be *recursively transfer lower semi-continuous* on  $\text{int } S^{L-1}$  with respect to  $D$  if, whenever  $q \in \text{int } S^{L-1}$  is not a competitive equilibrium price system, there exists some price system  $p^0 \in \text{int } S^{L-1}$  (possibly  $p^0 = q$ ) and a neighborhood  $\mathcal{V}_q$  such that (1) whenever  $p^0$  is upset by a price system in  $\text{int } S^{L-1} \setminus D$ , it is upset by a price system in  $D$ , and (2)  $p \cdot \hat{z}(\mathcal{V}_q) > 0$  for any  $p \in D$  that recursively upsets  $p^0$ .

Now we have the following theorem that fully characterizes the existence of competitive equilibrium in economies with possibly indivisible commodity spaces and discontinuous excess demand functions.

**Theorem 7.4.5 (The Existence Theorem I''')** Suppose an excess demand function  $\hat{z}(\cdot) : \text{int } S^{L-1} \rightarrow \mathbb{R}^L$  satisfies interior Walras' law:  $p \cdot \hat{z}(p) = 0$  for all  $p \in \text{int } S^{L-1}$ . Then there is a competitive price equilibrium  $p^* \in \text{int } S^{L-1}$  if and only if there exists a compact subset  $D \subseteq \text{int } S^{L-1}$  such that  $\hat{z}(\cdot)$  is recursively transfer lower semi-continuous on  $\text{int } S^{L-1}$  with respect to  $D$ .

**Proof.** *Sufficiency* ( $\Leftarrow$ ). The proof of sufficiency is essentially the same as that of Theorem 7.4.3 and we just outline the proof here. To show the existence of a competitive equilibrium on  $S^{L-1}$ , it suffices to show that there exists a competitive equilibrium  $p^*$  in  $D$  if it is recursively diagonal transfer lower semi-continuous on  $S^{L-1}$  with respect to  $D$ . Suppose, by way of contradiction, that there is no competitive equilibrium in  $D$ . Then, since  $\hat{z}$  is recursively diagonal transfer lower semi-continuous on  $S^{L-1}$  with respect to  $D$ , for each  $q \in D$ , there exists  $p^0$  and a neighborhood  $\mathcal{V}_q$  such that (1) whenever  $p^0$  is upset by a price system in  $S^{L-1} \setminus D$ , it is upset by a price system in  $D$  and (2)  $p \cdot \hat{z}(\mathcal{V}_q) > 0$  for any finite subset of price systems  $\{p^1, \dots, p^m\} \subset D$  with  $p^m = p$  and  $p \cdot \hat{z}(p^{m-1}) > 0$ ,  $p^{m-1} \cdot \hat{z}(p^{m-2}) > 0, \dots, p^1 \cdot \hat{z}(p^0) > 0$  for  $m \geq 1$ . Since there is no competitive equilibrium by the contrapositive hypothesis,  $p^0$  is not a competitive equilibrium and thus, by recursive diagonal transfer lower semi-continuity on  $S^{L-1}$  with respect to  $D$ , such a sequence of recursive securing price systems  $\{p^1, \dots, p^{m-1}, p\}$  exists for some  $m \geq 1$ .

Since  $D$  is compact and  $D \subseteq \bigcup_{q \in S^{L-1}} \mathcal{V}_q$ , there is a finite set  $\{q^1, \dots, q^T\} \subseteq D$  such that  $D \subseteq \bigcup_{i=1}^T \mathcal{V}_{q^i}$ . For each of such  $q^i$ , the corresponding initial deviation price system is denoted by  $p^{0i}$  so that  $p^i \cdot \hat{z}(\mathcal{V}_{q^i}) > 0$  whenever  $p^{0i}$  is recursively upset by  $p^i$  through any finite subset of securing price systems  $\{p^{1i}, \dots, p^{mi}\} \subset D$  with  $p^{mi} = p^i$ . Then, by the same argument as in the proof of Theorem 7.4.3, we will obtain that  $z^k$  is not in the union of  $\mathcal{V}_{q^1}, \mathcal{V}_{q^2}, \dots, \mathcal{V}_{q^k}$  for  $k = 1, 2, \dots, T$ . For  $k = T$ , we have  $p^T \notin \mathcal{V}_{q^1} \cup \mathcal{V}_{q^2} \dots \cup \mathcal{V}_{q^T}$  and so  $p^T \notin D \subseteq \bigcup_{i=1}^T \mathcal{V}_{q^i}$ , which contradicts that  $p^T$  is an upsetting price in  $D$ .

Thus, there exists a price system  $p^* \in S^{L-1}$  such that  $p \cdot \hat{z}(p^*) \leq 0$  for all  $p \in \text{int } S^{L-1}$ . We want to show that  $p^*$  in fact is a competitive price equilibrium. Note that  $\text{int } S^{L-1}$  is open and  $D$  is a compact subset of  $\text{int } S^{L-1}$ . One can always find a sequence of price vector  $\{q_n^l\} \subseteq \text{int } S^{L-1} \setminus D$  such that  $q_n^l \rightarrow p^l$  as  $n \rightarrow \infty$ , where  $p^l = (0, \dots, 0, 1, 0, \dots, 0)$  is the unit vector that has only one argument - the  $l$ th argument - with value 1 and others with value 0. Since  $p \cdot \hat{z}(q)$  is continuous in  $p$ , we have  $\hat{z}^l(p^*) \leq 0$  for  $l = 1, \dots, L$  and thus  $p^*$  is a competitive price equilibrium.

*Necessity* ( $\Rightarrow$ ). Suppose  $p^*$  is a competitive equilibrium. Let  $D = \{p^*\}$ . Then, the set  $D$  is clearly compact. Now, for any non competitive equilibrium  $q \in S^{L-1}$ , let  $p^0 = p^*$  and  $\mathcal{V}_q$  be a neighborhood of  $q$ . Since  $p \cdot \hat{z}(p^*) \leq 0$  for all  $p \in S^{L-1}$  and  $p^0 = p^*$  is a unique element in  $D$ , there is no other upsetting price  $p^1$  such that  $p^1 \cdot \hat{z}(p^0) > 0$ . Hence,



the game is recursively diagonal transfer continuous on  $S^{L-1}$  with respect to  $D$ . ■

Theorem 7.4.5 then strictly generalizes all the existing results on the existence of competitive equilibrium in economies with single-valued excess demand functions.

- **Conditions for Walras' Law to be true**

From the above theorems, Walras' Law is important to prove the existence of a competitive equilibrium. Under which conditions, is Walras' Law held?

When each consumer's budget constraint holds with equality:

$$px_i(p) = pw_i + \sum_{j=1}^J \theta_{ij} py_j(p)$$

for all  $i$ , we have

$$\begin{aligned} \sum_{i=1}^n px_i(p) &= \sum_{i=1}^n pw_i + \sum_{i=1}^n \sum_{j=1}^J \theta_{ij} py_j(p) \\ &= \sum_{i=1}^n pw_i + \sum_{j=1}^J py_j(p) \end{aligned}$$

which implies that

$$p \left[ \sum_{i=1}^n x_i(p) - \sum_{i=1}^n w_i - \sum_{j=1}^J y_j(p) \right] = 0 \quad (7.22)$$

so that

$$p \cdot \hat{z}(p) = 0 \quad (\text{Walras' Law}) \quad (7.23)$$

Thus, as long as the budget line holds with equality and aggregate excess demand is well-defined on the domain of prices, strong Walras' Law must hold.

The above existence theorems on competitive equilibrium are based on the assumptions that the aggregate excess demand correspondence is single-valued and satisfies the strong Walras's Law. The questions are under what conditions on economic environments a budget constraint holds with equality for all optimal consumption bundles, and the aggregate excess demand correspondence is single-valued or convex-valued? The following various types of monotonicities and convexities of preferences with the first one strongest and the last one weakest may be used to answer these questions.

- Types of monotonicity conditions

- (1) Strict monotonicity: For any two consumption market bundles  $(x \geq x')$  with  $x \neq x'$  implies  $x \succ_i x'$ .
- (2) Monotonicity: if  $x > x'$  implies that  $x \succ_i x'$ .
- (3) Local non-satiation: For any point  $x$  and any neighborhood,  $N(x)$ , there is  $x' \in N(x)$  such that  $x' \succ_i x$ .
- (4) Non-satiation: For any  $x$ , there exists  $x'$  such that  $x' \succ_i x$ .

**Remark 7.4.5** Monotonicity of preferences can be interpreted as individuals' desires for goods: the more, the better. Local non-satiation means individuals' desires are unlimited.

- Types of convexities

- (i) Strict convexity: For any  $x$  and  $x'$  with  $x \succ_i x'$  and  $x \neq x'$ ,  $x_\lambda \equiv \lambda x + (1 - \lambda)x' \succ_i x'$  for  $\lambda \in (0, 1)$ .

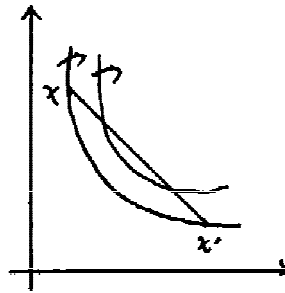


Figure 7.10: Strict convex indifference curves.

- (ii) Convexity: If  $x \succ_i x'$ , then  $x_\lambda = \lambda x + (1 - \lambda)x' \succ_i x'$  for  $\lambda \in (0, 1)$ .

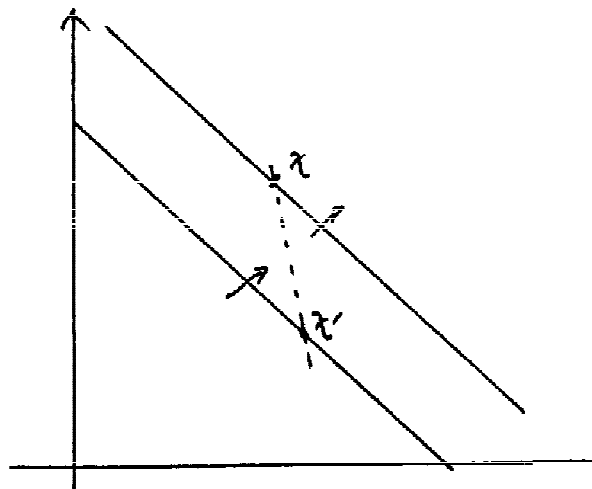


Figure 7.11: Linear indifference curves are convex, but not strict convex.

(iii) Weak convexity: If  $x \succsim_i x'$ , then  $x_\lambda \succsim_i x'$ .

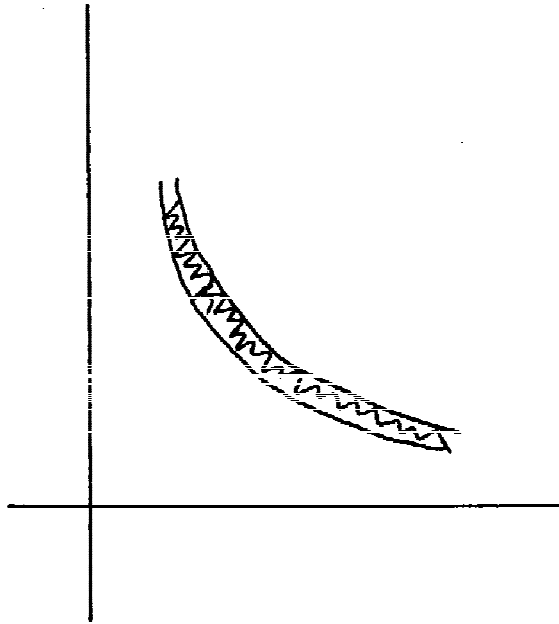


Figure 7.12: “Thick” indifference curves are weakly convex, but not convex.

The above conditions are ordered by increasing weakness, i.e., each of the above condition implies the next one, the converse may not be true by examining thick indifference curves and linear indifference curves, strictly convex indifference curves as showed in the above figures:

**Remark 7.4.6** The convexity of preferences implies that people want to diversify their consumptions, and thus, convexity can be viewed as the formal expression of basic measure of economic markets for diversification. Note that the strict convexity of  $\succsim_i$  implies the conventional diminishing marginal rates of substitution (MRS), and weak convexity of  $\succsim_i$  is equivalent to the quasi-concavity of utility function  $u_i$ . Also notice that the continuity of  $\succsim_i$  is a sufficient condition for the continuous utility representations, that is, it guarantees the existence of continuous utility function  $u_i(\cdot)$ .

**Remark 7.4.7** Under the convexity of preferences  $\succsim_i$ , non-satiation implies local non-satiation. Why? The proof is left to readers.

Now we are ready to answer under which conditions strong Walras's Law holds, a demand correspondence can be function, and convex-valued. The following propositions answer the questions.

**Proposition 7.4.1** *Under local non-satiation assumption, we have the budget constraint holds with equality, and thus the Walras's Law holds.*

**Proposition 7.4.2** *Under the strict convexity of  $\succsim_i$ ,  $x_i(p)$  becomes a (single-valued) function.*

**Proposition 7.4.3** *Under the weak convexity of preferences, demand correspondence  $x_i(p)$  is convex-valued.*

*Strict convexity of production set:* if  $y_j^1 \in Y_j$  and  $y_j^2 \in Y_j$ , then the convex combination  $\lambda y_j^1 + (1 - \lambda)y_j^2 \in \text{int}Y_j$  for all  $0 < \lambda < 1$ , where  $\text{int}Y_j$  denotes the interior points of  $Y_j$ .

The proof of the following proposition is based on the maximum theorem.

**Proposition 7.4.4** *If  $Y_j$  is compact (i.e., closed and bounded) and strictly convex, then the supply correspondence  $y_j(p)$  is a well defined single-valued and continuous function.*

Proof: By the maximum theorem, we know that  $y_j(p)$  is a non-empty valued upper hemi-continuous correspondence by the compactness of  $Y_j$  (by noting that  $0 \in Y_j$ ) for all  $p \in R_+^L$ .

Now we show it is single-valued. Suppose not.  $y_j^1$  and  $y_j^2$  are two profit maximizing production plans for  $p \in \mathbb{R}_+^L$ , and thus  $py_j^1 = py_j^2$ . Then, by the strict convexity of  $Y_j$ , we have  $\lambda y_j^1 + (1 - \lambda)y_j^2 \in \text{int}Y_j$  for all  $0 < \lambda < 1$ . Therefore, there exists some  $t > 1$  such that

$$t[\lambda y_j^1 + (1 - \lambda)y_j^2] \in \text{int}Y_j. \quad (7.24)$$

Then  $t[\lambda py_j^1 + (1 - \lambda)py_j^2] = tpy_j^1 > py_j^1$  which contradicts the fact that  $y_j^1$  is a profit maximizing production plan.

So  $y_j(p)$  is a single-valued function. Thus, by the upper hemi-continuity of  $y_j(p)$ , we know it is a single-valued and continuous function.

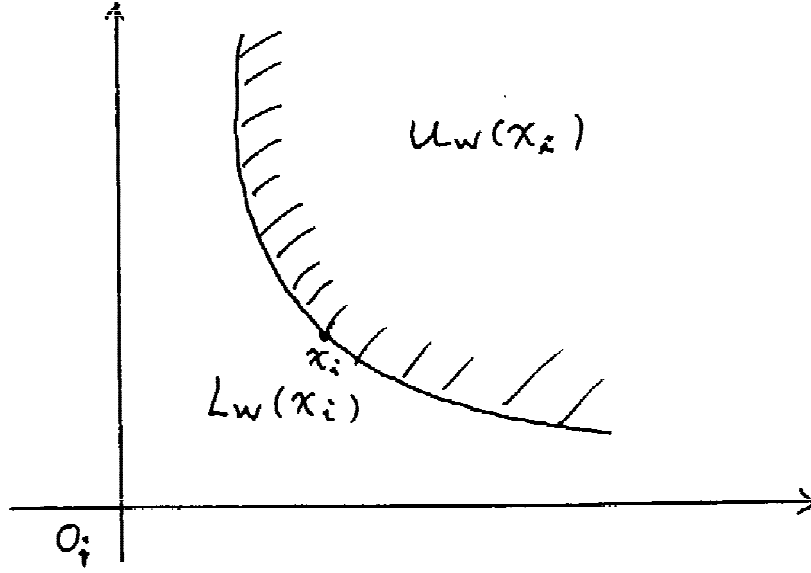


Figure 7.13: The upper contour set  $U_w(x_i)$  is given by all points above the indifference curve, and the lower contour set  $L_w(x_i)$  is given by all points below the indifference curve in the figure.

**Proposition 7.4.5** *If  $\succsim_i$  is continuous, strictly convex, locally non-satiated, and  $w_i > 0$ , then  $x_i(p)$  is a continuous single-valued function and satisfies the budget constraint with equality for all  $p \in \mathfrak{R}_{++}^L$ . Consequently, (strong) Walras's Law is satisfied for all  $p \in \mathfrak{R}_{++}^L$ .*

Proof. First note that, since  $w_i > 0$ , one can show that the budget constrained set  $B_i(p)$  is a continuous correspondence with non-empty and compact values and  $\succsim_i$  is continuous. Then, by the maximum theorem, we know the demand correspondence  $x_i(p)$  is upper hemi-continuous. Furthermore, by the strict convexity of preferences, it is single-valued and continuous. Finally, by local non-satiation, we know the budget constraint holds with equality, and thus strong Walras's Law is satisfied.

Then, from the above propositions, we can have the following existence theorem that provides sufficient conditions directly based on the fundamentals of the economy by applying the Existence Theorem  $I'''$  above.

**Theorem 7.4.6** *For a private ownership economy  $e = (\{X_i, w_i, \succsim_i\}, \{Y_j\}, \{\theta_{ij};\})$ , there exists a competitive equilibrium if the following conditions hold*

- (i)  $X_i \in \mathbb{R}_+^L$ ;
- (ii)  $\succsim_i$  are continuous, strictly convex (which guarantee the demand function is single valued) and strictly monotonic (which guarantees strong Walras' Law holds);
- (iii)  $w_i > 0$ ;
- (iv)  $Y_j$  are compact, strictly convex,  $0 \in Y_j$   $j = 1, 2, \dots, J$ .

Note that  $\succsim_i$  are continuous if the upper contour set  $U_w(x_i) \equiv \{x'_i \in X_i \text{ and } x'_i \succsim_i x_i\}$  and the lower contour set  $L_w(x_i) \equiv \{x'_i \in X_i \text{ and } x'_i \precsim_i x_i\}$  are closed.

Proof: By the assumption imposed, we know that  $x_i(p)$  and  $y_j(p)$  are continuous and single-valued. Thus the aggregate excess demand function is a continuous single-valued function and satisfies interior Walras' Law by monotonicity of preferences. Thus, we only need to show conditions (iv) and (v) in Theorem I''' are also satisfied. The bound in (iv) follows from the nonnegativity of demand (i.e., the fact that  $X_i = \mathbb{R}_+^L$ ) and bounded production sets, which implies that a consumer's total net supply to the market of any good  $l$  can be no greater the sum of his initial endowment and upper bound of production sets. Finally, we show that condition v is satisfied. As some prices go to zero, a consumer whose wealth tends to a strictly positive limit (note that, because  $p\hat{w} > 0$ , there must be at least one such consumer) and with strict monotonicity of preferences will demand an increasing large amount of some of the commodities whose prices go to zero. Hence, by Theorem I''', there is a competitive equilibrium. The proof is completed.

## Examples of Computing CE

**Example 7.4.1** Consider an exchange economy with two consumers and two goods with

$$\begin{aligned}
 w_1 &= (1, 0) & w_2 &= (0, 1) \\
 u_1(x_1) &= (x_1^1)^a (x_1^2)^{1-a} & 0 < a < 1 \\
 u_2(x_2) &= (x_2^1)^b (x_2^2)^{1-b} & 0 < b < 1
 \end{aligned} \tag{7.25}$$

Let  $p = \frac{p^2}{p^1}$

Consumer 1's problem is to solve

$$\max_{x_1} u_1(x_1) \tag{7.26}$$

subject to

$$x_1^1 + px_1^2 = 1 \quad (7.27)$$

Since utility functions are Cobb-Douglas types of functions, the solution is then given by

$$x_1^1(p) = \frac{a}{1} = a \quad (7.28)$$

$$x_1^2(p) = \frac{1-a}{p}. \quad (7.29)$$

Consumer 2's problem is to solve

$$\max_{x_2} u_2(x_2) \quad (7.30)$$

subject to

$$x_2^1 + px_2^2 = p. \quad (7.31)$$

The solution is given by

$$x_2^1(p) = \frac{b \cdot p}{1} = b \cdot p \quad (7.32)$$

$$x_2^2(p) = \frac{(1-b)p}{p} = (1-b). \quad (7.33)$$

Then, by the market clearing condition,

$$x_1^1(p) + x_2^1(p) = 1 \Rightarrow a + bp = 1 \quad (7.34)$$

and thus the competitive equilibrium is given by

$$p = \frac{p^2}{p^1} = \frac{1-a}{b}. \quad (7.35)$$

This is true because, by Walras's Law, for  $L = 2$ , it is enough to show only one market clearing.

**Remark 7.4.8** Since the Cobb-Douglas utility function is widely used as an example of utility functions that have nice properties such as strict monotonicity on  $\mathbb{R}_{++}^L$ , continuity, and strict quasi-concavity, it is useful to remember the functional form of the demand function derived from the Cobb-Douglas utility functions. It may be remarked that we can easily derive the demand function for the general function:  $u_i(x_i) = (x_i^1)^\alpha (x_i^2)^\beta$   $\alpha > 0, \beta > 0$  by the a suitable monotonic transformation. Indeed, by invariant to monotonic transformation of utility function, we can rewrite the utility function as

$$[(x_1^1)^\alpha (x_1^2)^\beta]^{\frac{1}{\alpha+\beta}} = (x_1^1)^{\frac{\alpha}{\alpha+\beta}} (x_1^2)^{\frac{\beta}{\alpha+\beta}} \quad (7.36)$$

so that we have

$$x_1^1(p) = \frac{\frac{\alpha}{\alpha+\beta}I}{p^1} \quad (7.37)$$

and

$$x_1^2(p) = \frac{\frac{\beta}{\alpha+\beta}I}{p^2} \quad (7.38)$$

when the budget line is given by

$$p^1 x_1^1 + p^2 x_1^2 = I. \quad (7.39)$$

### Example 7.4.2

$$\begin{aligned} n &= 2 & L &= 2 \\ w_1 &= (1, 0) & w_2 &= (0, 1) \\ u_1(x_1) &= (x_1^1)^a (x_1^2)^{1-a} & 0 < a < 1 \\ u_2(x_2) &= \min\{x_2^1, bx_2^2\} & \text{with } b > 0 \end{aligned} \quad (7.40)$$

For consumer 1, we have already obtained

$$x_1^1(p) = a, \quad x_1^2 = \frac{(1-a)}{p}. \quad (7.41)$$

For Consumer 2, his problem is to solve:

$$\max_{x_2} u_2(x_2) \quad (7.42)$$

s.t.

$$x_2^1 + px_2^2 = p. \quad (7.43)$$

At the optimal consumption, we have

$$x_2^1 = bx_2^2. \quad (7.44)$$

By substituting the solution into the budget equation, we have  $bx_2^2 + px_2^2 = p$  and thus

$$x_2^2(p) = \frac{p}{b+p} \text{ and } x_2^1(p) = \frac{bp}{b+p}.$$

Then, by  $x_1^1(p) + x_2^1(p) = 1$ , we have

$$a + \frac{bp}{b+p} = 1 \quad (7.45)$$

or

$$(1-a)(b+p) = bp \quad (7.46)$$



so that

$$(a + b - 1)p = b(1 - a) \quad (7.47)$$

Thus,

$$p^* = \frac{b(1 - a)}{a + b - 1}.$$

To make  $p^*$  be a competitive equilibrium price, we need to assume  $a + b > 1$ .

Note that there is no competitive equilibrium when  $a + b \leq 1$  although the aggregate excess demand function is continuous on the domain. What causes the nonexistence of competitive equilibrium and which condition cannot be dropped?

## 7.4.2 The Existence of CE for Aggregate Excess Demand Correspondences

When preferences and/or production sets are not strictly convex, the demand correspondence and/or supply correspondence may not be single-valued, and consequently the aggregate excess demand correspondence may not be single-valued. As a result, one cannot use the above existence results to argue the existence of competitive equilibrium. Nevertheless, by using the KKM lemma, we can still prove the existence of competitive equilibrium when the aggregate excess demand correspondence satisfies certain conditions.

**Theorem 7.4.7 (The Existence Theorem II)** *For a private ownership economy  $e = (\{X_i, w_i, \succsim_i\}, \{Y_j\}, \{\theta_{ij}\})$ , if  $\hat{z}(p)$  is an non-empty convex and compact-valued upper hemi-continuous correspondence and satisfies weak Walras' Law, then there exists a competitive equilibrium, that is, there is a price vector  $p^* \in S$  such that*

$$\hat{z}(p^*) \cap \{-\mathcal{R}_+^L\} \neq \emptyset. \quad (7.48)$$

Proof. Define a correspondence  $F : S \rightarrow 2^S$  by,

$$F(q) = \{p \in S : q\hat{z} \leq 0 \text{ for some } \hat{z} \in \hat{z}(p)\}.$$

Since  $\hat{z}(\cdot)$  is upper hemi-continuous,  $F(q)$  is closed for each  $q \in S$ . We now prove  $F$  is FS-convex. Suppose, by way of contradiction, that there are some  $q_1, \dots, q_m \in S$  and some convex combination  $q_\lambda = \sum_{t=1}^m \lambda_t q_t$  such that  $q_\lambda \notin \cup_{t=1}^m F(q_t)$ . Then,  $q_\lambda \notin F(q_t)$  for all  $t = 1, \dots, m$ . Thus, for all  $\hat{z} \in \hat{z}(q_\lambda)$ , we have  $q_t \hat{z} > 0$  for  $t = 1, \dots, m$ . Hence,

$\sum_{t=1}^m \lambda_t q_t \hat{z} = q_\lambda \hat{z} > 0$  which contradicts the fact that  $\hat{z}$  satisfies Walras' Law. So  $F$  must be FS-convex. Therefore, by KKM lemma, we have

$$\cap_{q \in S} F(q) \neq \emptyset.$$

Then there exists a  $p^* \in S$  such that  $p^* \in \cap_{q \in S} F(q)$ , i.e.,  $p^* \in F(q)$  for all  $q \in S$ . Thus, for each  $q \in S$ , there is  $\hat{z}_q \in \hat{z}(p^*)$  such that

$$q \hat{z}_q \leq 0.$$

We now prove  $\hat{z}(p^*) \cap \{-\mathfrak{R}_+^L\} \neq \emptyset$ . Suppose not. Since  $\hat{z}(p^*)$  is convex and compact and  $-\mathfrak{R}_+^L$  is convex and closed, by the Separating Hyperplane theorem, there exists a vector  $q \in \mathfrak{R}^L$  and some  $c \in \mathfrak{R}^L$  such that

$$q \cdot (-\mathfrak{R}_+^L) < c < q \hat{z}(p^*)$$

Since  $(-\mathfrak{R}_+^L)$  is a cone, we must have  $c > 0$  and  $q \cdot (-\mathfrak{R}_+^L) \leq 0$ . Thus,  $q \in \mathfrak{R}_+^L$  and  $q \hat{z}(p^*) > 0$  for all  $q$ , a contradiction. The proof is completed.

**Remark 7.4.9** The last part of the proof can be also shown by applying the following result: Let  $K$  be a compact convex set. Then  $K \cap \{-\mathfrak{R}_+^L\} \neq \emptyset$  if and only if for any  $p \in \mathfrak{R}_+^L$ , there exists  $z \in K$  such that  $pz \leq 0$ . The proof of this result can be found, for example, in the book of K. Border (1985, p. 13).

Similarly, we have the following existence theorem that provides sufficient conditions directly based on economic environments by applying the Existence Theorem II above.

**Theorem 7.4.8** *For a private ownership economy  $e = (\{X_i, w_i, \succsim_i\}), \{Y_j\}, \{\theta_{ij}\}$ , there exists a competitive equilibrium if the following conditions hold*

- (i)  $X_i \in \mathfrak{R}_+^L$ ;
- (ii)  $\succsim_i$  are continuous, weakly convex and strictly monotone;
- (iii)  $\hat{w} > 0$ ;
- (iv)  $Y_j$  are closed, convex, and  $0 \in Y_j$ ;  $j = 1, 2, \dots, J$ .

### 7.4.3 The Existence of CE for General Production Economies

For a general private ownership production economy,

$$e = (\{X_i, \succsim_i, w_i\}, \{Y_j\}, \{\theta_{ij}\}) \quad (7.49)$$

recall that a competitive equilibrium consists of a feasible allocation  $(x^*, y^*)$  and a price vector  $p^* \in R_+^L$  such that

$$(i) \ x_i^* \in D_i(p^*) \equiv x_i(p^*) \quad (\text{utility maximization, } i = 1, 2, \dots, n)$$

$$(ii) \ y_j^* \in S_j(p^*) \equiv y_j(p^*) \quad (\text{profit maximization, } j = 1, 2, \dots, J)$$

We now state the following existence theorem for general production economies without proof. Since the proof is very complicated, we refer readers to the proof that can be found in Debreu (1959) who used the existence theorem on equilibrium of the abstract economy on Section 7.7.

**Theorem 7.4.9 (Existence Theorem III, Debreu, 1959)** *A competitive equilibrium for the private-ownership economy  $e$  exists if the following conditions are satisfied:*

- (1)  $X_i$  is closed, convex and bounded from below;
- (2)  $\succsim_i$  are non-satiated;
- (3)  $\succsim_i$  are continuous;
- (4)  $\succsim_i$  are convex;
- (5)  $w_i \in \text{int}X_i$ ;
- (6)  $0 \in Y_j$  (possibility of inaction);
- (7)  $Y_j$  are closed and convex (continuity and no IRS)
- (8)  $Y_j \cap \{-Y_j\} = \{0\}$  (Irreversibility)
- (9)  $\{-\mathfrak{R}_+^L\} \subseteq Y_j$  (free disposal)

## 7.5 The Uniqueness of Competitive Equilibria

So now we know that a Walrasian equilibrium will exist under some regularity conditions such as continuity, monotonicity, and convexity of preferences and/or production possibility sets. We worry next about the other extreme possibility; for a given economy there

after many equilibria so that we do not know what is the outcome of the market process. We can easily give examples in which there are multiple competitive equilibrium price vectors. When is there only one normalized price vector that clears all markets?

The free goods case is not of great interest here, so we will rule it out by means of the desirability assumption so that every equilibrium price of each good must be strictly positive. We want to also assume the continuous differentiability of the aggregate excess demand function. The reason is fairly clear; if indifference curves have kinks in them, we can find whole ranges of prices that are market equilibria. Not only are the equilibria not unique, they are not even locally unique.

Thus, we answer this question for only considering the case of  $p^* > 0$  and  $\hat{z}(p)$  is differentiable.

**Theorem 7.5.1** *Suppose all goods are desirable and gross substitute for all prices (i.e.,  $\frac{\partial z^h(p)}{\partial p^l} > 0$  for  $l \neq h$ ). If  $p^*$  is a competitive equilibrium price vector and Walras' Law holds, then it is the unique competitive equilibrium price vector.*

Proof: By the desirability,  $p^* > 0$ . Suppose  $p$  is another competitive equilibrium price vector that is not proportional to  $p^*$ . Let  $m = \max \frac{p^l}{p^{*l}} = \frac{p^k}{p^{*k}}$  for some  $k$ . By homogeneity and the definition of competitive equilibrium, we know that  $\hat{z}(p^*) = \hat{z}(mp^*) = 0$ . We know that  $m = \frac{p^k}{p^{*k}} \geq \frac{p^l}{p^{*l}}$  for all  $l = 1, \dots, L$  and  $m > \frac{p^h}{p^{*h}}$  for some  $h$ . Then we have  $mp^{*l} \geq p^l$  for all  $l$  and  $mp^{*h} > p^h$  for some  $h$ . Thus, when the price of good  $k$  is fixed, the prices of the other goods are down. We must have the demand for good  $k$  down by the gross substitutes. Hence, we have  $\hat{z}^k(p) < 0$ , a contradiction.

When the aggregate demand function satisfies the Weak Axiom of Revealed Preference (WARP) and Walras' Law holds, competitive equilibrium is also unique.

**The Weak Axiom of Revealed Preference (WARP) of the aggregate excess demand function:** If  $p\hat{z}(p) \geq p\hat{z}(p')$ , then  $p'\hat{z}(p) > p'\hat{z}(p')$  for all  $p, p' \in \mathbb{R}_+^L$ .

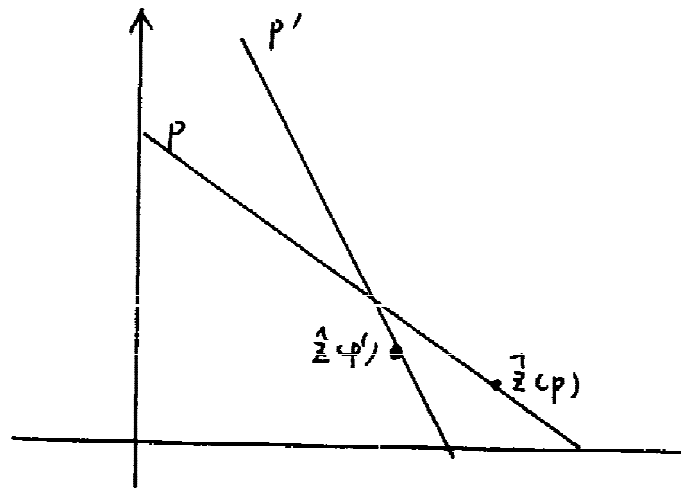


Figure 7.14: The figure shows an aggregate demand function satisfies WARP.

WARP implies that, if  $\hat{z}(p')$  could have been bought at  $p$  where  $\hat{z}(p)$  was bought (so that  $\hat{z}(p) \succ \hat{z}(p')$  since  $\hat{z}(p)$  is the optimal choice), then at price  $p'$ ,  $\hat{z}(p)$  is outside the budget constraint (otherwise it contradicts to the fact that  $\hat{z}(p)$  is the optimal choice).

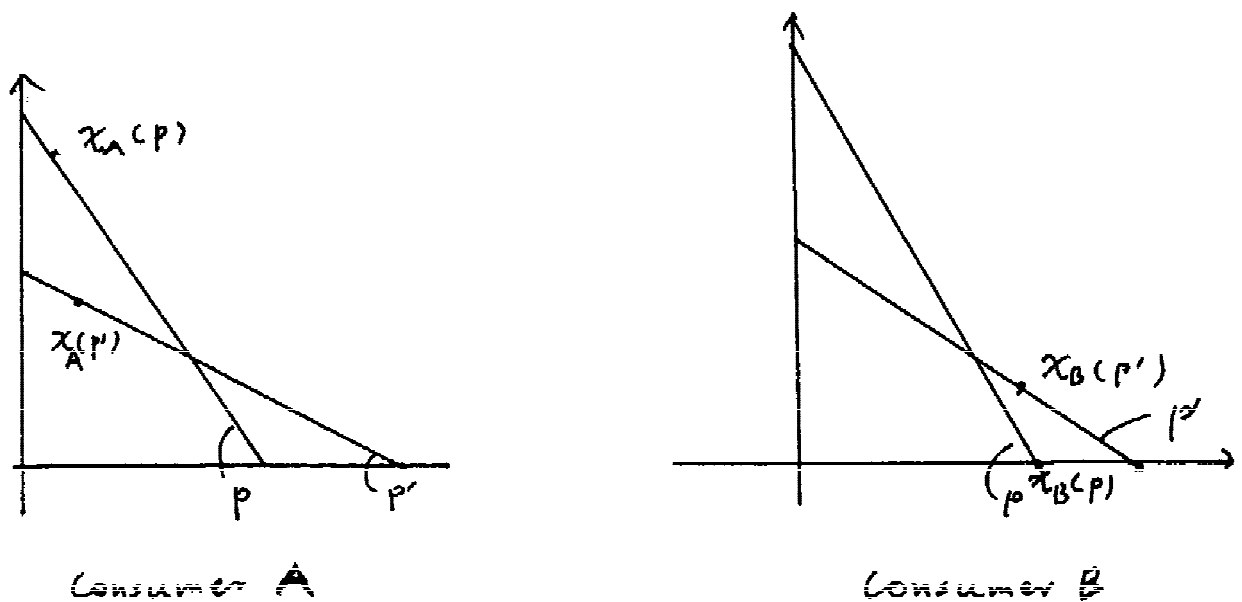


Figure 7.15: Both individual demand functions satisfy WARP.

Even if the individual demand functions satisfy WARP, the aggregate excess demand function does not necessarily satisfy the WARP in aggregate.

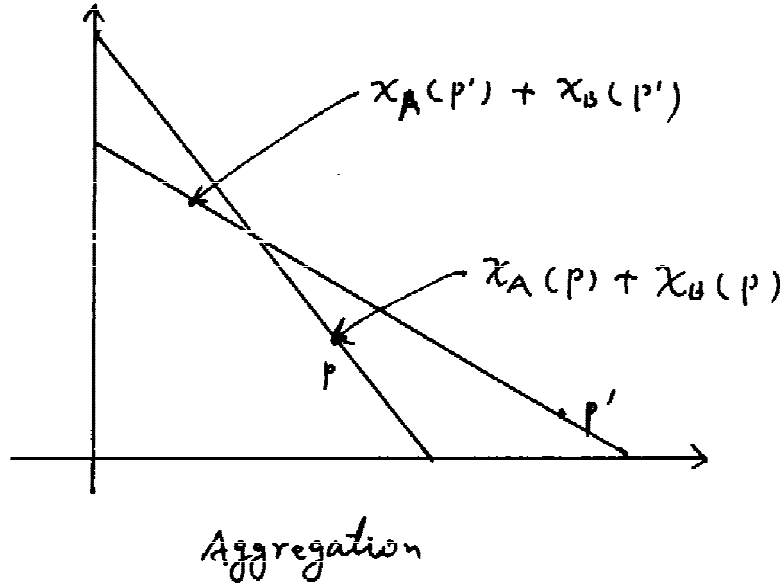


Figure 7.16: The aggregate excess demand function does not satisfy WARP.

The WARP is a weak restriction than the continuous concave utility function. However, the restriction on the aggregate excess demand function is not as weak as it may be seen. Even though two individuals satisfy the individual WARP, the aggregate excess demand function may not satisfy the aggregate WARP as shown in Figures 7.15 and 7.16.

**Lemma 7.5.1** *Under the assumptions of strong Walras' Law and WARP, we have  $p^* \hat{z}(p) > 0$  for all  $p \neq kp^*$  where  $p^*$  is a competitive equilibrium.*

Proof: Suppose  $p^*$  is a competitive equilibrium:

$$\hat{z}(p^*) \leq 0. \quad (7.50)$$

Also, by Walras' Law,  $p \hat{z}(p) = 0$ . So we have  $p \hat{z}(p) \geq p \hat{z}(p^*)$ . Then, by WARP,  $p^* \hat{z}(p) > p^* \hat{z}(p^*) = 0 \Rightarrow p^* \hat{z}(p) > 0$  for all  $p \neq kp^*$ .

**Theorem 7.5.2** *Under the assumptions of Walras' Law and WARP in aggregate excess demand function, the competitive equilibrium is unique.*

Proof: By Lemma 7.5.1, for any  $p \neq kp^*$ ,  $p^* \hat{z}(p) > 0$  which means at least for some  $l$ ,  $\hat{z}^l > 0$ .

## 7.6 Stability of Competitive Equilibrium

The concept of competitive equilibrium is a stationary concept. But, it has given no guarantee that the economy will actually operate at the equilibrium point. Or an economy may deviate from an equilibrium due to economic shocks. What forces exist that might tend to move prices to a market-clearing price? This is a topic about the stability on the price adjustment mechanism in a competitive equilibrium.

A paradoxical relationship between the idea of competition and price adjustment is that: If all agents take prices as given, how can prices move? Who is left to adjust prices?

To solve this paradox, one introduces a “Walrasian auctioneer” whose sole function is to seek for the market clearing prices. The Walrasian auctioneer is supposed to call the prices and change the price mechanically responding to the aggregate excess demand till the market clears. Such a process is called Tâtonnement adjustment process.

**Tâtonnement Adjustment Process** is defined, according to the laws of demand and supply, by

$$\frac{dp^l}{dt} = G^l(\hat{z}^l(p)) \quad l = 1, \dots, L \quad (7.51)$$

where  $G^l$  is a sign-preserving function of  $\hat{z}^l(p)$ , i.e.,  $G^l(x) > 0$  if  $x > 0$ ,  $G^l(x) = 0$  if  $x = 0$ , and  $G^l(x) < 0$  if  $x < 0$ . The above equation implies that when the aggregate excess demand is positive, we have a shortage and thus price should go up by the laws of demand and supply.

As a special case of  $G^l$ ,  $G^l$  can be an identical mapping such that

$$\dot{p}^l = \hat{z}^l(p) \quad (7.52)$$

$$\dot{p} = \hat{z}(p). \quad (7.53)$$

Under Walras' Law,

$$\begin{aligned} \frac{d}{dt}(p'p) &= \frac{d}{dt} \left[ \sum_{l=1}^L (p^l)^2 \right] = 2 \sum_{l=1}^L (p^l) \cdot \frac{dp^l}{dt} \\ &= 2p'\dot{p} = p\hat{z}(p) = 0 \end{aligned}$$

which means that the sum of squares of the prices remain constant as the price adjusts. This is another price normalization. The path of the prices are restricted on the surface of a  $k$ -dimensional sphere.

Examples of price dynamics in Figure 7.17. The first and third figures show a stable equilibrium, the second and fourth figures show a unique unstable equilibrium. Formally, we have

**Definition 7.6.1** An equilibrium price  $p^*$  is globally stable if

- (i)  $p^*$  is the unique competitive equilibrium,
- (ii) for all  $p_o$  there exists a unique price path  $p = \phi(t, p_o)$  for  $0 \leq t < \infty$  such that  $\lim_{t \rightarrow \infty} \phi(t, p_o) = p^*$ .

**Definition 7.6.2** An equilibrium price  $p^*$  is locally stable if there is  $\delta > 0$  and a unique price path  $p = \phi(t, p_o)$  such that  $\lim_{t \rightarrow \infty} \phi(t, p_o) = p^*$  whenever  $|p^* - p_o| < \delta$ .

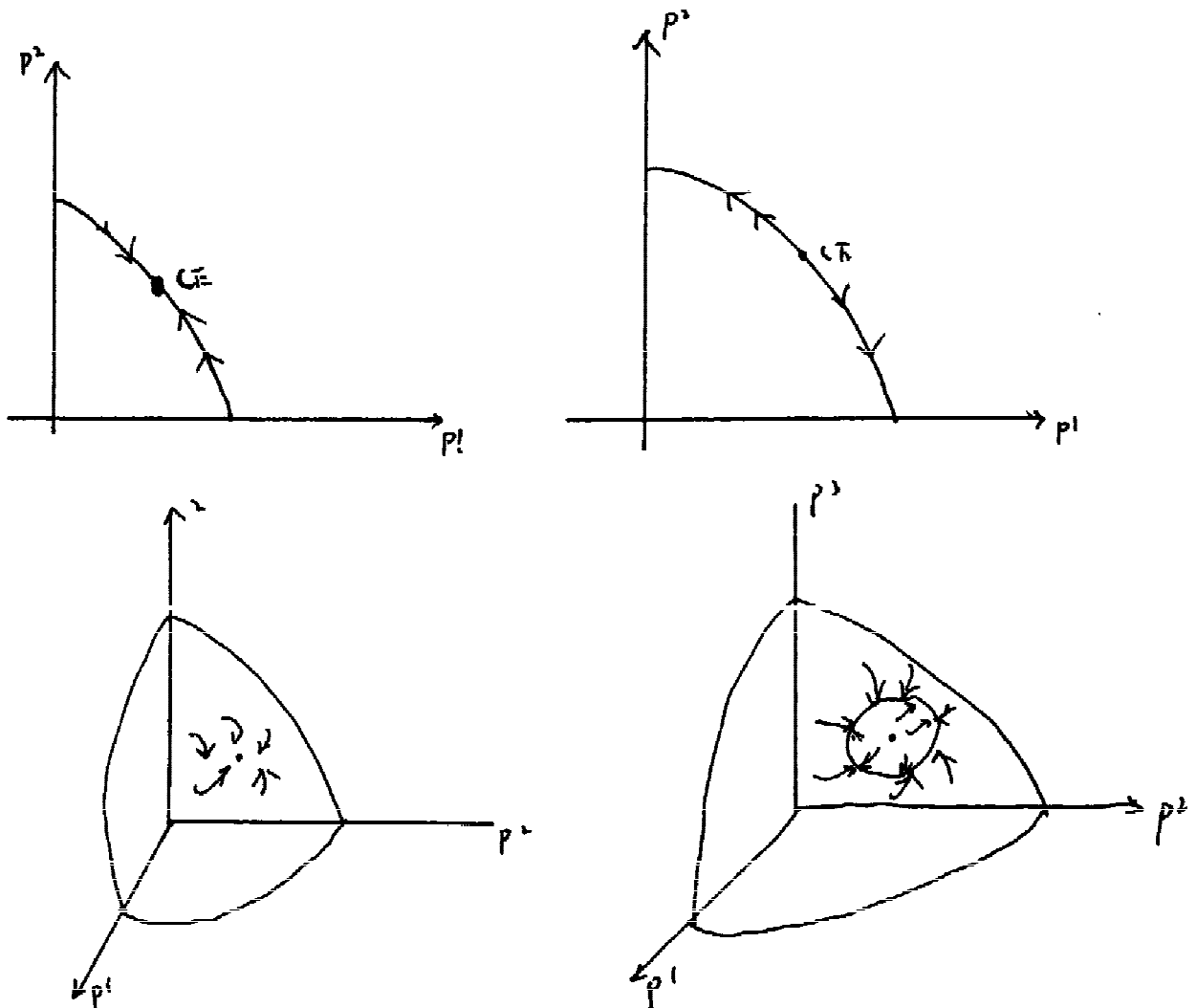




Figure 7.17: The first and third figures show the CEs are stable, and the second and fourth show they are not stable.

The local stability of a competitive equilibrium can be easily obtained from the standard result on the local stability of a differential equation.

**Theorem 7.6.1** *A competitive equilibrium price  $p^*$  is locally stable if the Jacobean matrix defined by*

$$A = \left[ \frac{\partial \hat{z}^l(p^*)}{\partial p_k} \right]$$

*has all negative characteristic roots.*

The global stability result can be obtained by Liapunov Theorem.

**Liapunov's function:** For a differential equation system  $\dot{x} = f(x)$  with  $f(x^*) = 0$ , a function  $V$  is said to be a Liapunov's function for the differential equation system if

- (1) there is a unique  $x^*$  such that  $V(x^*) = 0$ ;
- (2)  $V(x) > 0$  for all  $x \neq x^*$ ;
- (3)  $\frac{dV(x)}{dt} < 0$ .

**Theorem 7.6.2 (Liapunov's Theorem)** *If there exists a Liapunov's function for  $\dot{x} = f(x)$ , the unique stationary point  $x^*$  is globally stable.*

Debreu (1974) has shown essentially that any continuous function which satisfies the Walras' Law is an aggregate demand function for some economy. Thus, utility maximization places no restriction on aggregate behavior. Thus, to get global stability, one has to make some additional assumptions such as gross substitutability and the Weak Axiom of Revealed Preference (WARP) for aggregate excess demand functions.

**Lemma 7.6.1** *Under the assumptions of Walras' Law and gross substitutability, we have  $p^* \hat{z}(p) > 0$  for all  $p \neq kp^*$  where  $p^* > 0$  is a competitive equilibrium.*

Proof. The proof of this lemma is complicated. We illustrate the proof only for the case of two commodities by aid of the figure. The general proof of this lemma can be seen in Arrow and Hahn, 1971, p. 224.

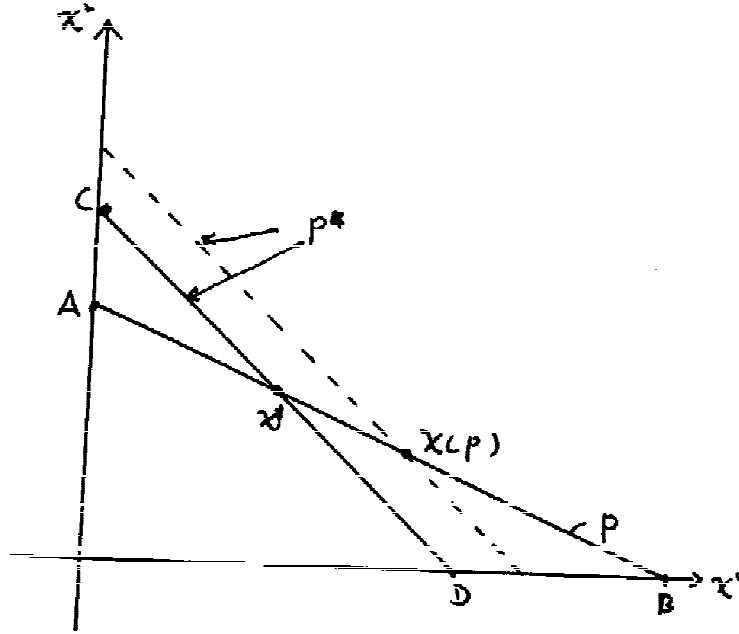


Figure 7.18: An illustration of the proof of Lemma 7.6.1.

Let  $p^*$  be an equilibrium price vector and  $\hat{x}^* = \hat{x}(p^*)$  be the aggregate equilibrium demand. Let  $p \neq \alpha p^*$  for any  $\alpha > 0$ . Then we know  $p$  is not an equilibrium price vector by the uniqueness of the competitive equilibrium under the gross substitutability. Since  $p\hat{x}(p) = p\hat{w} = p\hat{x}(p^*)$  by Walras' Law and  $\hat{w} = \hat{x}(p^*)$ , then the aggregate demand  $\hat{x}(p)$  is on the line  $AB$  which passes through the point  $\hat{x}^*$  and whose slope is given by  $p$ . Let  $CD$  be the line which passes through the point  $\hat{x}^*$  and whose slope is  $p^*$ . We assume that  $p^{*1}/p^{*2} > p^1/p^2$  without loss of generality. Then  $p^{*1}/p^1 > p^{*2}/p^2 \equiv \mu$ . Thus,  $p^{*1} > \mu p^1$  and  $p^{*2} = \mu p^2$ . Therefore, we have  $\hat{x}^2(p^*) > \hat{x}^2(\mu p)$  by the gross substitutability. But Walras' Law implies that  $\mu p\hat{x}(\mu p) = \mu p\hat{x}(p^*)$  so that we must have  $\hat{x}^1(\mu p) > \hat{x}^1(p^*)$ . Using the homogeneity assumption, we get  $\hat{x}^1(p) = \hat{x}^1(\mu p) > \hat{x}^1(p^*) = \hat{x}^{*1}$  and  $\hat{x}^2(p) = \hat{x}^2(\mu p) < \hat{x}^2(p^*) = \hat{x}^{*2}$ . Hence the point  $\hat{x}(p)$  must lie to the right of the point  $\hat{x}^*$  in the figure. Now draw a line parallel to  $CD$  passing through the point  $\hat{x}(p)$ . We see that  $p^*\hat{x}(p) > p^*\hat{x}^*$  and thus  $p^*\hat{z}(p) > 0$ . The proof is completed.

Now we are ready to prove the following theorem on the global stability of a competitive equilibrium.

**Theorem 7.6.3 (Arrow-Block-Hurwicz)** *Under the assumptions of Walras' Law, if*

$\hat{z}(p)$  satisfies either gross substitutability, or WARP, then the competitive equilibrium price is globally stable.

Proof: By the gross substitutability or WARP of the aggregate excess demand function, the competitive  $p^*$  is unique. We now show it is globally stable. Define a Liapounov's function by

$$V(p) = \sum_{l=1}^L (p^l(t) - p^{*l})^2 = (p - p^*) \cdot (p - p^*) \quad (7.54)$$

By the assumption, the competitive equilibrium price  $p^*$  is unique. Also, since

$$\begin{aligned} \frac{dV}{dt} &= 2 \sum_{l=1}^L (p(t)^l - p^{*l}) \frac{dp^l(t)}{dt} \\ &= 2 \sum_{l=1}^L (p_{(t)}^l - p^{*l}) \hat{z}^l(p) \\ &= 2[p\hat{z}(p) - p^*\hat{z}(p)] \\ &= -2p^*\hat{z}(p) < 0 \end{aligned}$$

by Walras' Law and Lemma 7.5.1 and 7.6.1 for  $p \neq p^*$ , we know  $\dot{p} = \hat{z}(p)$  is globally stable by Leapounov's theorem.

The above theorem establishes the global stability of a competitive equilibrium under Walras' Law, homogeneity, and gross substitutability/WARP. It is natural to ask how far we can relax these assumptions, in particular gross substitutability. Since many goods in reality are complementary goods, can the global stability of a competitive equilibrium also hold for complementary goods? Scarf (1961) has constructed examples that show that a competitive equilibrium may not be globally stable in the presence of complementary goods.

**Example 7.6.1 (Scarf's Counterexample on Global Instability)** Consider a pure exchange economy with three consumers and three commodities ( $n=3$ ,  $L=3$ ). Suppose consumers' endowments are given by  $w_1 = (1, 0, 0)$ ,  $w_2 = (0, 1, 0)$   $w_3 = (0, 0, 1)$  and their utility functions are given by

$$\begin{aligned} u_1(x_1) &= \min\{x_1^1, x_1^2\} \\ u_2(x_2) &= \min\{x_2^2, x_2^3\} \\ u_3(x_3) &= \min\{x_3^1, x_3^3\} \end{aligned}$$

so that they have  $L$ -shaped indifference curves. Then, the aggregate excess demand function is given by

$$\hat{z}^1(p) = -\frac{p^2}{p^1 + p^2} + \frac{p^3}{p^1 + p^3} \quad (7.55)$$

$$\hat{z}^2(p) = -\frac{p^3}{p^2 + p^3} + \frac{p^1}{p^1 + p^2} \quad (7.56)$$

$$\hat{z}^3(p) = -\frac{p^1}{p^1 + p^3} + \frac{p^2}{p^2 + p^3}, \quad (7.57)$$

from which the only possible competitive equilibrium price system is given by  $p^{*1} = p^{*2} = p^{*3} = 1$ .

Then, the dynamic adjustment equation is given by

$$\dot{p} = \hat{z}(p).$$

We know that  $\|p(t)\| = \text{constant}$  for all  $t$  by Walras' Law. Now we want to show that  $\prod_{l=1}^3 p^l(t) = \text{constant}$  for all  $t$ . Indeed,

$$\begin{aligned} \frac{d}{dt} \left( \prod_{l=1}^3 p^l(t) \right) &= \dot{p}^1 p^2 p^3 + \dot{p}^2 p^1 p^3 + \dot{p}^3 p^1 p^2 \\ &= \hat{z}^1 p^2 p^3 + \hat{z}^2 p^1 p^3 + \hat{z}^3 p^1 p^2 = 0. \end{aligned}$$

Now, we show that the dynamic process is not globally stable. First choose the initial prices  $p^l(0)$  such that  $\sum_{l=1}^3 [p^l(0)]^2 = 3$  and  $\prod_{l=1}^3 p^l(0) \neq 1$ . Then,  $\sum_{l=1}^3 [p^l(t)]^2 = 3$  and  $\prod_{l=1}^3 p^l(t) \neq 1$  for all  $t$ . Since  $\sum_{l=1}^3 [p^l(t)]^2 = 3$  and the only possible equilibrium prices are  $p^{*1} = p^{*2} = p^{*3} = 1$ , the solution of the above system of differential equations cannot converge to the equilibrium price  $p^* = (1, 1, 1)$ .

In this example, we may note the following facts. (i) there is no substitution effect, (ii) the indifference curve is not strictly convex, and (iii) the difference curve has a kink and hence is not differentiable. Scarf (1961) also provided the examples of instability in which the substitution effect is present for the case of Giffen's goods. Thus, Scarf's examples indicate that instability may occur in a wide variety of cases.

## 7.7 Abstract Economy

The abstract economy defined by Debreu (Econometrica, 1952) generalizes the notion of  $N$ -person Nash noncooperative game in that a player's strategy set depends on the

strategy choices of all the other players and can be used to prove the existence of competitive equilibrium since the market mechanism can be regarded as an abstract economy as shown by Arrow and Debreu (Econometrica 1954). Debreu (1952) proved the existence of equilibrium in abstract economies with finitely many agents and finite dimensional strategy spaces by assuming the existence of continuous utility functions. Since Debreu's seminal work on abstract economies, many existence results have been given in the literature. Shafer and Sonnenschein (J. of Mathematical Economics, 1975) extended Debreu's results to abstract economies without ordered preferences.

### 7.7.1 Equilibrium in Abstract Economy

Let  $N$  be the set of agents which is any countable or uncountable set. Each agent  $i$  chooses a strategy  $x_i$  in a set  $X_i$  of  $\mathbb{R}^L$ . Denote by  $X$  the (Cartesian) product  $\prod_{j \in N} X_j$  and  $X_{-i}$  the product  $\prod_{j \in N \setminus \{i\}} X_j$ . Denote by  $x$  and  $x_{-i}$  an element of  $X$  and  $X_{-i}$ . Each agent  $i$  has a payoff (utility) function  $u_i : X \rightarrow \mathbb{R}$ . Note that agent  $i$ 's utility is not only dependent on his own choice, but also dependent on the choice of the others. Given  $x_{-i}$  (the strategies of others), the choice of the  $i$ -th agent is restricted to a non-empty feasible choice set  $F_i(x_{-i}) \subset X_i$ , the  $i$ -th agent chooses  $x_i \in F_i(x_{-i})$  so as to maximize  $u_i(x_{-i}, x_i)$  over  $F_i(x_{-i})$ .

An abstract economy (or called generalized game)  $\Gamma = (X_i, F_i, u_i)_{i \in N}$  is defined as a family of ordered triples  $(X_i, F_i, u_i)$ .

**Definition 7.7.1** A vector  $x^* \in X$  is said to be an *equilibrium of an abstract economy* if  $\forall i \in N$

- (i)  $x_i^* \in F_i(x_{-i}^*)$  and
- (ii)  $x_i^*$  maximizes  $u_i(x_{-i}^*, x_i)$  over  $F_i(x_{-i}^*)$ .

If  $F_i(x_{-i}) \equiv X_i$ ,  $\forall i \in N$ , the abstract economy reduces to the conventional game  $\Gamma = (X_i, u_i)$  and the equilibrium is called a *Nash equilibrium*.

**Theorem 7.7.1 (Arrow-Debreu)** Let  $X$  be a non-empty compact convex subset of  $\mathbb{R}^{nL}$ . Suppose that

- i) the correspondence  $F: X \rightarrow 2^X$  is a continuous correspondence with non-empty compact and convex values,
- ii)  $u_i: X \times X \rightarrow \mathbb{R}$  is continuous,
- iii)  $u_i: X \times X \rightarrow \mathbb{R}$  is either quasi-concave in  $x_i$  or it has a unique maximum on  $F_i(x_{-i})$  for all  $x_{-i} \in X_{-i}$ .

Then  $\Gamma$  has an equilibrium.

Proof. For each  $i \in N$ , define the maximizing correspondence

$$M_i(x_{-i}) = \{x_i \in F_i(x_{-i}) : u_i(x_{-i}, x_i) \geq u_i(x_{-i}, z_i), \forall z_i \in F_i(x_{-i})\}.$$

Then, the correspondence  $M_i: X_{-i} \rightarrow 2^{X_i}$  is non-empty compact and convex valued because  $u_i$  is continuous in  $x$  and quasi-concave in  $x_i$  and  $F_i(x_{-i})$  is non-empty convex compact-valued. Also, by the Maximum Theorem,  $M_i$  is an upper hemi-continuous correspondence. Therefore the correspondence

$$M(x) = \prod_{i \in N} M_i(x_{-i})$$

is an upper hemi-continuous correspondence with non-empty convex compact-values. Thus, by Kakutani's Fixed Point Theorem, there exists  $x^* \in X$  such that  $x^* \in M(x^*)$  and  $x^*$  is an equilibrium in the generalized game. Q.E.D

A competitive market mechanism can be regarded as an abstract economy. For simplicity, consider an exchange economy  $e = (X_i, u_i, w_i)_{i \in N}$ . Define an abstract economy  $\Gamma = (Z, F_i, u_i)_{i \in N+1}$  as follows. Let

$$\begin{aligned} Z_i &= X_i \quad i = 1, \dots, n \\ Z_{n+1} &= S^{L-1} \\ F_i(x_i, p) &= \{x_i \in X_i : px_i \leq pw_i\} \quad i = 1, \dots, n \\ F_{n+1} &= S^{L-1} \\ u_{n+1}(p, x) &= \sum_{i=1}^n p(x_i - w_i) \end{aligned} \tag{7.58}$$

Here  $N + 1$ 's sole purpose is to change prices, and thus can be regarded as a Walrasian auctioneer. Then, we verify that the economy  $e$  has a competitive equilibrium if the abstract economy defined above has an equilibrium by noting that  $\sum_{i=1}^n x_i \leq \sum_{i=1}^n w_i$  at the equilibrium of the abstract equilibrium.

### 7.7.2 The Existence of Equilibrium for General Preferences

The above theorem on the existence of an equilibrium in abstract economy has assumed the preference relation is an ordering and can be represented by a utility function. In this subsection, we consider the existence of equilibrium in an abstract economy where individuals' preferences  $\succsim$  may be non total or not-transitive. Define agent  $i$ 's preference correspondence  $P_i : X \rightarrow 2^{X_i}$  by

$$P_i(x) = \{y_i \in X_i : (y_i, x_{-i}) \succsim_i (x_i, x_{-i})\}$$

We call  $\Gamma = (X_i, F_i, P_i)_{i \in N}$  an abstract economy.

A *generalized game (or an abstract economy)*  $\Gamma = (X_i, F_i, P_i)_{i \in N}$  is defined as a family of ordered triples  $(X_i, F_i, P_i)$ . An *equilibrium* for  $\Gamma$  is an  $x^* \in X$  such that  $x^* \in F(x^*)$  and  $P_i(x^*) \cap F_i(x^*) = \emptyset$  for each  $i \in N$ .

Shafer and Sonnenschein (1975) proved the following theorem that generalizes the above theorem to an abstract economy with non-complete/non-transitive preferences.

**Theorem 7.7.2 (Shafer-Sonnenschein )** *Let  $\Gamma = (X_i, F_i, P_i)_{i \in N}$  be an abstract economy satisfying the following conditions for each  $i \in N$ :*

- (i)  $X_i$  is a non-empty, compact, and convex subset in  $\mathbb{R}^{l_i}$ ,
- (ii)  $F_i : X \rightarrow 2^{X_i}$  is a continuous correspondence with non-empty, compact, and convex values,
- (iii)  $P_i$  has open graph,
- (iv)  $x_i \notin \text{con } P_i(x)$  for all  $x \in Z$ .

*Then  $\Gamma$  has an equilibrium.*

This theorem requires the preferences have open graph. Tian (International Journal of Game Theory, 1992) proved the following theorem that is more general and generalizes the results of Debreu (1952), Shafer and Sonnenschein (1975) by relaxing the openness of graphs or lower sections of preference correspondences. Before proceeding to the theorem, we state some technical lemmas which were due to Micheal (1956, Propositions 2.5, 2.6 and Theorem 3.1''').

**Lemma 7.7.1** *Let  $X \subset \mathbb{R}^M$  and  $Y \subset \mathbb{R}^K$  be two convex subsets and  $\phi : X \rightarrow 2^Y$ ,  $\psi : X \rightarrow 2^Y$  be correspondences such that*

- (i)  *$\phi$  is lower hemi-continuous, convex valued, and has open upper sections,*
- (ii)  *$\psi$  is lower hemi-continuous,*
- (iii) *for all  $x \in X$ ,  $\phi(x) \cap \psi(x) \neq \emptyset$ .*

*Then the correspondence  $\theta : X \rightarrow 2^Y$  defined by  $\theta(x) = \phi(x) \cap \psi(x)$  is lower hemi-continuous.*

**Lemma 7.7.2** *Let  $X \subset \mathbb{R}^M$  and  $Y \subset \mathbb{R}^K$  be two convex subsets, and let  $\phi : X \rightarrow 2^Y$  be lower hemi-continuous. Then the correspondence  $\psi : X \rightarrow 2^Y$  defined by  $\psi(x) = \text{con } \phi(x)$  is lower hemi-continuous.*

**Lemma 7.7.3** *Let  $X \subset \mathbb{R}^M$  and  $Y \subset \mathbb{R}^K$  be two convex subsets. Suppose  $F : X \rightarrow 2^Y$  is a lower hemi-continuous correspondence with non-empty and convex values. Then there exists a continuous function  $f : X \rightarrow Y$  such that  $f(x) \in F(x)$  for all  $x \in X$ .*

**Theorem 7.7.3 (Tian)** *Let  $\Gamma = (X_i, F_i, P_i)_{i \in N}$  be a generalized game satisfying for each  $i \in N$ :*

- (i)  *$X_i$  is a non-empty, compact, and convex subset in  $\mathbb{R}^{l_i}$ ,*
- (ii)  *$F_i$  is a continuous correspondence, and  $F_i(x)$  is non-empty, compact, and convex for all  $x \in X$ ,*
- (iii)  *$P_i$  is lower hemi-continuous and has open upper sections,*
- (iv)  *$x_i \notin \text{con } P_i(x)$  for all  $x \in F$ .*

*Then  $\Gamma$  has an equilibrium.*

Proof. For each  $i \in N$ , define a correspondence  $A_i : X \rightarrow 2^{X_i}$  by  $A_i(x) = F_i(x) \cap \text{con } P_i(x)$ . Let  $U_i = \{x \in X : A_i(x) \neq \emptyset\}$ . Since  $F_i$  and  $P_i$  are lower hemi-continuous in  $X$ , so are they in  $U_i$ . Then, by Lemma 7.7.2,  $\text{con } P_i$  is lower hemi-continuous in  $U_i$ . Also since  $P_i$  has open upper sections in  $X$ , so does  $\text{con } P_i$  in  $X$  and thus  $\text{con } P_i$  in  $U_i$ . Further,  $F_i(x) \cap \text{con } P_i(x) \neq \emptyset$  for all  $x \in U_i$ . Hence, by Lemma 7.7.1, the correspondence  $A_i|_{U_i} :$



$U_i \rightarrow 2^{X_i}$  is lower hemi-continuous in  $U_i$  and for all  $x \in U_i$ ,  $F(x)$  is non-empty and convex. Also  $X_i$  is finite dimensional. Hence, by Lemma 7.7.3, there exists a continuous function  $f_i : U_i \rightarrow X_i$  such that  $f_i(x) \in A_i(x)$  for all  $x \in U_i$ . Note that  $U_i$  is open since  $A_i$  is lower hemi-continuous. Define a correspondence  $G_i : X \rightarrow 2^{X_i}$  by

$$G_i(x) = \begin{cases} \{f_i(x)\} & \text{if } x \in U_i \\ F_i(x) & \text{otherwise} \end{cases}. \quad (7.59)$$

Then  $G_i$  is upper hemi-continuous. Thus the correspondence  $G : X \rightarrow 2^X$  defined by  $G(x) = \prod_{i \in N} G_i(x)$  is upper hemi-continuous and for all  $x \in X$ ,  $G(x)$  is non-empty, closed, and convex. Hence, by Kakutani's Fixed Point Theorem, there exists a point  $x^* \in X$  such that  $x^* \in G(x^*)$ . Note that for each  $i \in N$ , if  $x^* \in U_i$ , then  $x_i^* = f_i(x^*) \in A(x^*) \subset \text{con } P_i(x^*)$ , a contradiction to (iv). Hence,  $x^* \notin U_i$  and thus for all  $i \in N$ ,  $x_i^* \in F_i(x^*)$  and  $F_i(x^*) \cap \text{con } P_i(x^*) = \emptyset$  which implies  $F_i(x^*) \cap P_i(x^*) = \emptyset$ . Thus  $\Gamma$  has an equilibrium. ■

Note that a correspondence  $P$  has open graph implies that it has upper and lower open sections; a correspondence  $P$  has lower open sections implies  $P$  is lower hemi-continuous. Thus, the above theorem is indeed weaker.

## Reference

- Arrow, K., H. D. Block, and L. Hurwicz, "On the Stability of the Competitive Equilibrium, II," *Econometrica*, 27 (1959), 82-109.
- Arrow, K and G. Debreu, "Existence of Equilibrium for a Competitive Economy," *Econometrica*, 1954, 22.
- Arrow, K., and L. Hurwicz, "On the Stability of the Competitive Equilibrium, I," *Econometrica*, 26 (1958), 522-552.
- Arrow, K. J. and Hahn, F. H. (1971), *General Competitive Analysis*, (San Francisco, Holden Day).
- Baye, M. J. Zhou, and Tian, G. "Characterizations of the Existence of Equilibria in Games with Discontinuous and Nonquasiconcave Payoffs," *Review of Economic Studies*, 60 (1993), 935-948.

- Border, K. C., *Fixed Point Theorems with Applications to Economics and Game Theory*, Cambridge: Cambridge University Press, 1985.
- Debreu, G., "A Social equilibrium existence theorem," *Proceedings of the National Academy of Sciences of the U. S. A.* 38 (1952) 386-393.
- Debreu, G., *Theory of Value*, (Wiley, New York), 1959.
- Debreu, G. "Excess Demand Functions," *Journal of Mathematical Economics*, 1 (1974), 15-22.
- Hildenbrand, W., A. P. Kirman, *Equilibrium Analysis: Variations on the Themes by Edgeworth and Walras*, North-Holland: New York, 1988.
- Jehle, G. A., and P. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, 1998, Chapter 7.
- Fan, K., "Some Properties of Convex Sets Related to Fixed Point Theorem," *Mathematics Annals*, 266 (1984), 519-537.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 7.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995, Chapters 15, 17.
- Michael, E., "Continuous selections I," *Annals of Mathematics* 63 (1956), 361-382.
- Scarf, H., "Some Examples of Global Instability of the Competitive Equilibrium," *International Economic Review*, 1 (1960), 157-172.
- Shafer, W. and H. Sonnenschein, Equilibrium in abstract economies without ordered preferences, *Journal of Mathematical Economics* 2 (1975) 345-348.
- Sonnenschein, H., *Demand Theory without Transitive Preferences, with Application to the Theory of Competitive Equilibrium*, Preferences, Utility, and Demand, Edited by J. S. Chipman, L. Hurwicz, M. K. Richter, and H. Sonnenschein, Harcourt Brace Jovanovich, New York, New York, 1971.

- Takayama, A. *Mathematical Economics*, the second edition, Cambridge: Cambridge University Press, 1985, Chapters 1-3.
- Tian, G. "On the Existence of Equilibria in Generalized Games," *International Journal of Game Theory*, 20 (1992), pp.247-254.
- Tian, G. "Existence of Equilibrium in Abstract Economies with Discontinuous Payoffs and Non-Compact Choice Spaces," *Journal of Mathematical Economics*, 21 (1992), pp. 379-388.
- Tian, G. (2009): Existence of Equilibria in Games with Arbitrary Strategy Spaces and Payoffs: A Full Characterization, mimeo.
- Tian, G. (2013): "On the Existence of Price Equilibrium in Economies with Excess Demand Functions", mimeo.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 17-18, 21.
- Walras, L., *Eléments of d'Économie Politique Pure*. Lausanne: Gorbaz, 1874. [Translated as *Elements of Pure Economics*. Homewood, Ill: Irwin, 1954]

# Chapter 8

## Normative Theory of Equilibrium: Its Welfare Properties

### 8.1 Introduction

In the preceding chapter, we studied the conditions which would guarantee the existence, uniqueness, and stability of competitive equilibrium. In this chapter, we will study the welfare properties of competitive equilibrium.

Economists are interested not only in describing the world of the market economy, but also in evaluating it. Does a competitive market do a “good” job in allocating resources? Adam Smith’s “invisible hand” says that market economy is efficient. Then in what sense and under what conditions is a market efficient? The concept of efficiency is related to a concern with the well-being of those in the economy. The normative analysis can not only help us understand what advantage the market mechanism has, it can also help us to evaluate an economic system used in the real world, as well as help us understand why China and East European countries want to change their economic institutions.

The term of economic efficiency consists of three requirements:

- (1) Exchange efficiency: goods are traded efficiently so that no further mutual beneficial trade can be obtained.
- (2) Production efficiency: there is no waste of resources in producing goods.
- (3) Product mix efficiency, i.e., the mix of goods produced by the economy

reflects the preferences of those in the economy.

## 8.2 Pareto Efficiency of Allocation

When economists talk about the efficiency of allocation resources, it means Pareto efficiency. It provides a minimum criterion for efficiency of using resources.

Under the market institution, we want to know what is the relationship between a market equilibrium and Pareto efficiency. There are two basic questions: (1) If a market (not necessarily competitive) equilibrium exists, is it Pareto efficient? (2) Can any Pareto efficient allocation be obtained through the market mechanism by redistributing endowments?

The concept of Pareto efficiency is not just to study the efficiency of a market economy, but it also can be used to study the efficiency of any economic system. First let us define the notion of Pareto improvement.

**Definition 8.2.1 (Pareto Improvement:)** An allocation can have a Pareto improvement if it is feasible and there is another feasible allocation such that one person would be better off and all other persons are not worse off.

**Definition 8.2.2 (Pareto Efficiency, also called Pareto Optimality:)** An allocation is Pareto efficient or Pareto optimal (in short, P.O) if there is no Pareto improvement, that is, if it is feasible and there is no other feasible allocation such that one person would be better off and all other persons are not worse off.

More precisely, for exchange economies, a feasible allocation  $x$  is *P.O.* if there is no other allocation  $x'$  such that

$$(i) \sum_{i=1}^n x'_i \leq \sum_{i=1}^n w_i$$

$$(ii) x'_i \succsim_i x_i \text{ for all } i \text{ and } x'_k \succ_k x_k \text{ for some } k = 1, \dots, n.$$

For production economy,  $(x, y)$  is Pareto optimal if and only if:

$$(1) \hat{x} \leq \hat{y} + \hat{w}$$

(2) there is no feasible allocation  $(x', y')$  s.t.

$$\begin{aligned} x'_i &\succ x_i \text{ for all } i \\ x'_k &\succ_k x_k \text{ for some } k \end{aligned}$$

A weaker concept about economic efficiency is the so-called weak Pareto efficiency.

**Definition 8.2.3 (Weak Pareto Efficiency)** An allocation is *weakly Pareto efficient* if it is feasible and there is no other feasible allocation such that all persons are better off.

**Remark 8.2.1** Some textbooks such as Varian (1992) has used weak Pareto optimality as the definition of Pareto optimality. Under which conditions are they equivalent? It is clear that Pareto efficiency implies weak Pareto efficiency. But the converse may not be true. However, under the continuity and strict monotonicity of preferences, the converse is true.

**Proposition 8.2.1** *Under the continuity and strict monotonicity of preferences, weak Pareto efficiency implies Pareto efficiency.*

Proof. Suppose  $x$  is weakly Pareto efficient but not Pareto efficient. Then, there exists a feasible allocation  $x'$  such that  $x'_i \succ_i x_i$  for all  $i$  and  $x'_k \succ_k x_k$  for some  $k$ .

Define  $\bar{x}$  as follows

$$\begin{aligned} \bar{x}_k &= (1 - \theta)x'_k \\ \bar{x}_i &= x'_i + \frac{\theta}{n-1}x'_k \text{ for } i \neq k \end{aligned}$$

Then we have

$$\bar{x}_k + \sum_{i \neq k} \bar{x}_i = (1 - \theta)x'_k + \sum_{i \neq k} (x'_i + \frac{\theta}{n-1}x'_k) = \sum_{i=1}^n x'_i \quad (8.1)$$

which means  $\bar{x}$  is feasible. Furthermore, by the continuity of preferences,  $\bar{x}_k = (1 - \theta)x'_k \succ x_k$  when  $\theta$  is sufficiently close to zero, and  $\bar{x}_i \succ_i x_i$  for all  $i \neq k$  by the strict monotonicity of preferences. This contradicts the fact that  $x$  is weakly Pareto optimal. Thus, we must have every weak Pareto efficient allocation is Pareto efficient under the monotonicity and continuity of preferences. ■

**Remark 8.2.2** The trick of taking a little from one person and then equally distribute it to the others will make every one better off is a useful one. We will use the same trick to prove the second theorem of welfare economics.

**Remark 8.2.3** The above proposition also depends on an implicit assumption that the goods under consideration are all private goods. Tian (Economics Letters, 1988) showed, by example that, if goods are public goods, we may not have such equivalence between Pareto efficiency and weak Pareto efficiency under the monotonicity and continuity of preferences.

The set of Pareto efficient allocations can be shown with the Edgeworth Box. Every point in the Edgeworth Box is attainable. A is the starting point. Is “A” a weak Pareto optimal? No. The point C, for example in the shaded area is better off to both persons. Is the point “B” a Pareto optimal? YES. Actually, all tangent points are Pareto efficient points where at least some agent does not have incentives to trade. The locus of all Pareto efficient is called the contract curve.

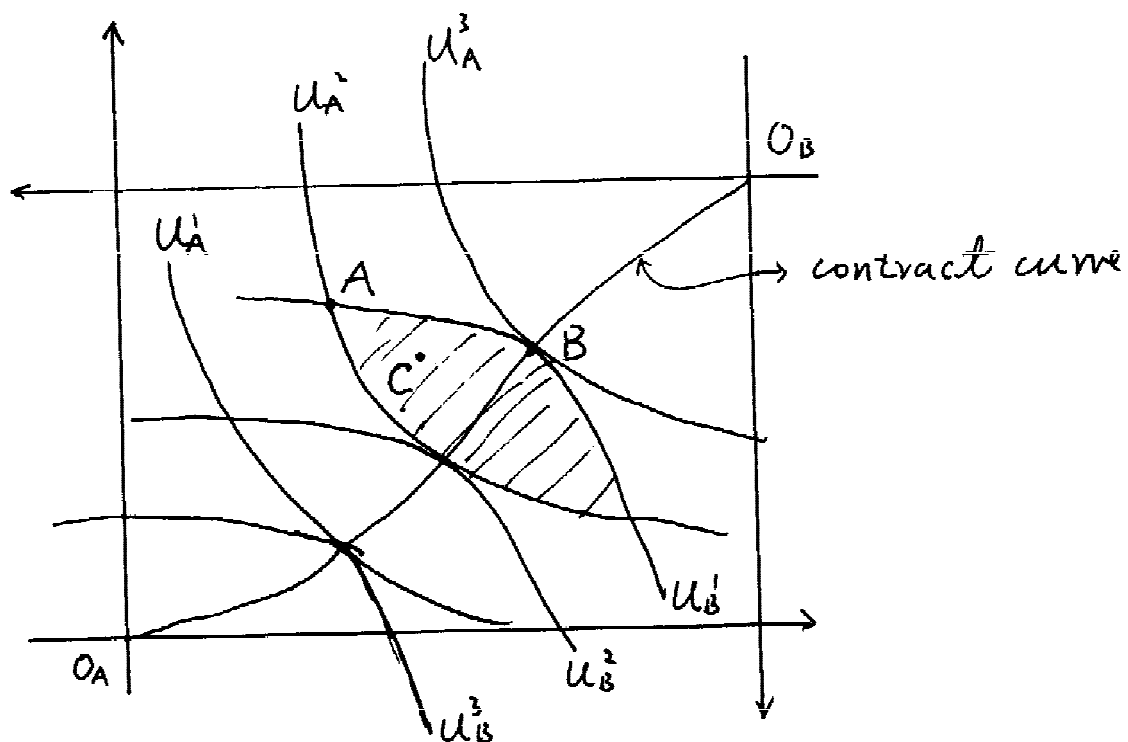


Figure 8.1: The set of Pareto efficient allocations is given by the contract curve.

**Remark 8.2.4** *Equity* and *Pareto efficiency* are two different concepts. The points like  $O_A$  or  $O_B$  are Pareto optimal points, but these are extremely unequal. To make an allocation be relatively equitable and also efficient, government needs to implement some institutional arrangements such as tax, subsidy to balance between equity and efficiency, but this is a value judgement and policy issue.

We will show that, when Pareto efficient points are given by tangent points of two persons' indifference curves, it should satisfy the following conditions:

$$MRS_{x^1x^2}^A = MRS_{x^1x^2}^B$$

$$x_A + x_B = \hat{w}.$$

When indifference curves of two agents are never tangent, we have the following cases.

Case 1. For linear indifference curves with non-zero (negative) slope, indifference curves of agents may not be tangent. How can we find the set of Pareto efficient allocations in this case? We can do it by comparing the steepness of indifference curves.

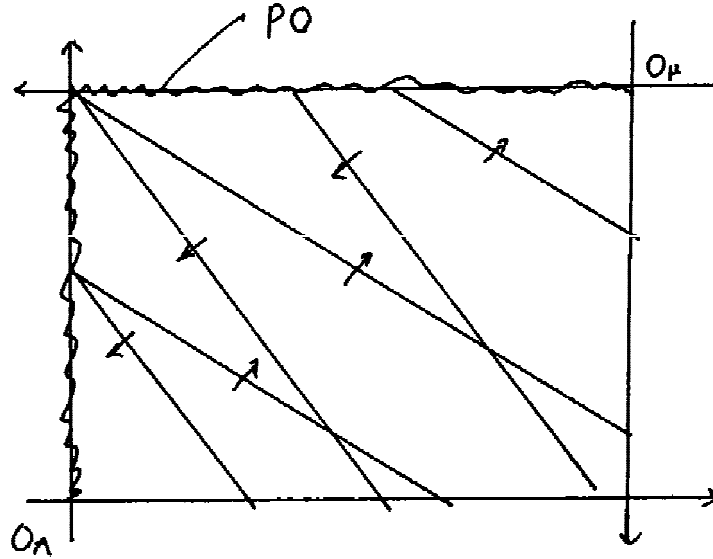


Figure 8.2: The set of Pareto efficient allocations is given by the upper and left edges of the box when indifference curves are linear and Person B's indifference curves are steeper.

$$MRS_{x^1x^2}^A < MRS_{x^1x^2}^B \quad (8.2)$$



In this case, when indifference curves for  $B$  is given, say, by Line  $AA$ , then  $K$  is a Pareto efficient point. When indifferent curves for  $A$  is given, say, by Line  $BB$ , then  $P$  is a Pareto efficient point. Contract curve is then given by the upper and left edge of the box. As a usual method of finding all possible Pareto efficient points, whether or not difference curves have tangent points, taken one person's any indifference as given, find another person's utility maximizing point. The maximizing point must be Pareto efficient point.

Case 2. Suppose that indifference curves are given by

$$u_A(x_A) = x_A^2$$

and

$$u_B(x_B) = x_B^1$$

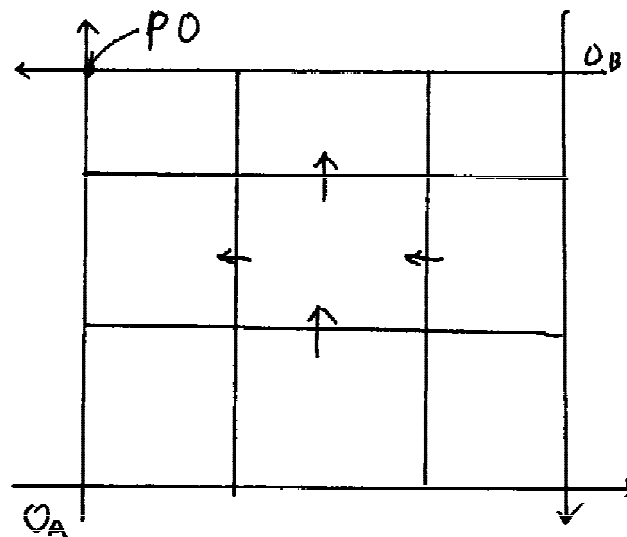


Figure 8.3: The only Pareto efficient allocation point is given by the upper and left corner of the box when individuals only care about one commodity.

Then, only Pareto efficient is the left upper corner point. But the set of weakly Pareto efficient is given by the upper and left edge of the box. Notice that utility functions in this example are continuous and monotonic, but a weak Pareto efficient allocation may not be Pareto efficient. This example shows that the strict monotonicity cannot be replaced by the monotonicity for the equivalence of Pareto efficiency and weak efficiency.

Case 3. Now suppose that indifference curves are perfect complementary. Then, utility functions are given by

$$u_A(x_A) = \min\{ax_A^1, bx_A^2\}$$

and

$$u_B(x_B) = \min\{cx_B^1, dx_B^2\}$$

A special case is the one where  $a = b = c = d = 1$ .

Then, the set of Pareto optimal allocations is the area given by points between two  $45^\circ$  lines.

Case 4. One person's indifference curves are "thick." In this case, an weak Pareto efficient allocation may not be Pareto efficient.

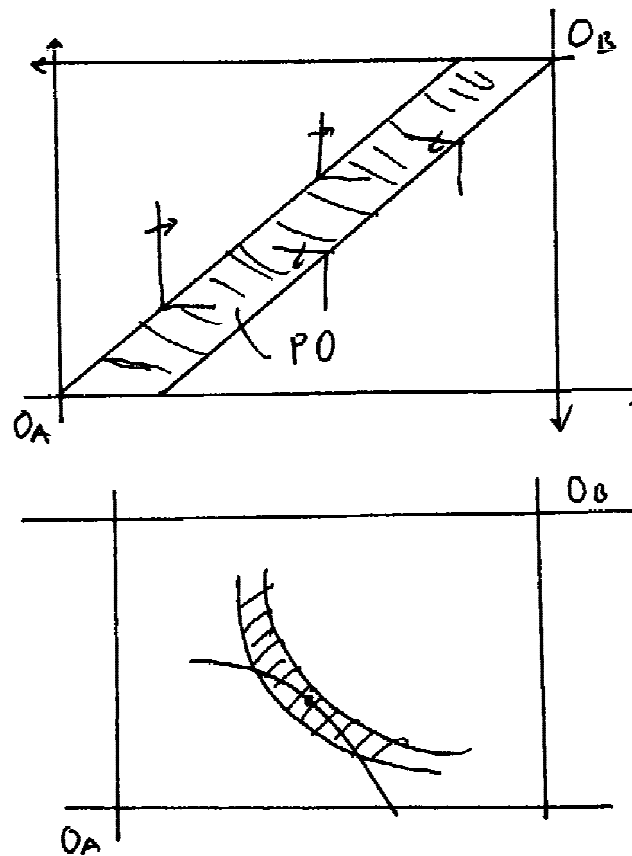


Figure 8.4: The first figure shows that the contract curve may be the "thick" when indifference curves are perfect complementary. The second figure shows that a weak Pareto efficient allocation may not Pareto efficient when indifference curves are "thick."

## 8.3 The First Fundamental Theorem of Welfare Economics

There is a well-known theorem, in fact, one of the most important theorems in economics, which characterizes a desirable nature of the competitive market institution. It claims that every competitive equilibrium allocation is Pareto efficient. A remarkable part of this theorem is that the theorem requires few assumptions, much fewer than those for the existence of competitive equilibrium. Some implicit assumptions in this section are that preferences are orderings, complete information goods are divisible, and there are no public goods, or externalities.

**Theorem 8.3.1 (The First Fundamental Theorem of Welfare Economics)** *If  $(x, y, p)$  is a competitive equilibrium, then  $(x, y)$  is weakly Pareto efficient, and further under local non-satiation, it is Pareto efficient.*

Proof: Suppose  $(x, y)$  is not weakly Pareto optimal, then there exists another feasible allocation  $(x', y')$  such that  $x'_i \succ_i x_i$  for all  $i$ . Thus, we must have  $px'_i > pw_i + \sum_{j=1}^J \theta_{ij}py_j$  for all  $i$ . Therefore, by summation, we have

$$\sum_{i=1}^n px'_i > \sum_{i=1}^n pw_i + \sum_{j=1}^J py_j.$$

Since  $py_j \geq py'_j$  for all  $y'_j \in Y_j$  by profit maximization, we have

$$\sum_{i=1}^n px'_i > \sum_{i=1}^n pw_i + \sum_{j=1}^J py'_j. \quad (8.3)$$

or

$$p\left[\sum_{i=1}^n x'_i - \sum_{i=1}^n w_i - \sum_{j=1}^J y'_j\right] > 0, \quad (8.4)$$

which contradicts the fact that  $(x', y')$  is feasible.

To show Pareto efficiency, suppose  $(x, y)$  is not Pareto optimal. Then there exists another feasible allocation  $(x', y')$  such that  $x'_i \succ_i x_i$  for all  $i$  and  $x'_k \succ_k x_k$  for some  $k$ . Thus we have, by local non-satiation,

$$px'_i \geq pw_i + \sum_{j=1}^J \theta_{ij}py_j \quad \forall i$$

and by  $x'_k \succ_k x_k$  why? readers can prove this,

$$px'_k > pw_k + \sum_{j=1}^J \theta_{kj} py_j$$

and thus

$$\sum_{i=1}^n px'_i > \sum_{i=1}^n pw_i + \sum_{j=1}^J py_j \geq \sum_{i=1}^n pw_i + \sum_{j=1}^J py'_j. \quad (8.5)$$

Again, it contradicts the fact that  $(x', y')$  is feasible. ■

**Remark 8.3.1** If the local non-satiation condition is not satisfied, a competitive equilibrium allocation  $x$  may not be Pareto optimal, say, for the case of thick indifference curves.

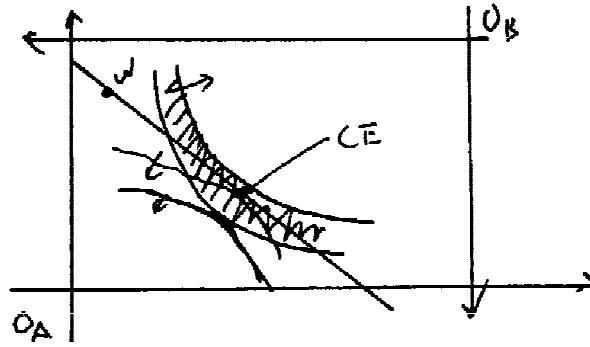


Figure 8.5: A CE allocation may not be Pareto efficient when the local non-satiation condition is not satisfied.

**Remark 8.3.2** Note that neither convexity of preferences nor convexity of production set is assumed in the theorem. The conditions required for Pareto efficiency of competitive equilibrium is much weaker than the conditions for the existence of competitive equilibrium.

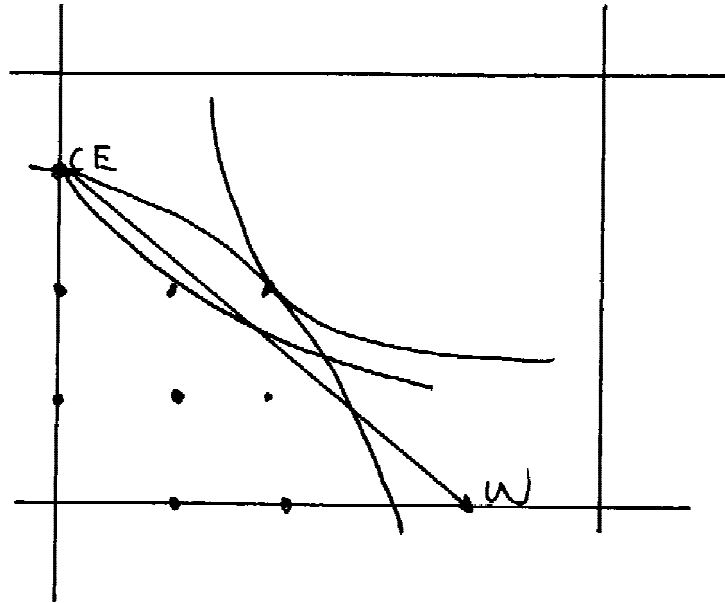


Figure 8.6: A CE allocation may not be Pareto efficient when goods are indivisible.

**Remark 8.3.3** If goods are indivisible, then a competitive equilibrium allocation  $x$  may not be Pareto optimal.

The point  $x$  is a competitive equilibrium allocation, but not Pareto optimal since  $x'$  is preferred by person 1. Hence, the divisibility condition cannot be dropped for the First Fundamental Theorem of Welfare Theorem to be true.

## 8.4 Calculations of Pareto Optimum by First-Order Conditions

The first-order conditions for Pareto efficiency are a bit harder to formulate. However, the following trick is very useful.

### 8.4.1 Exchange Economies

**Proposition 8.4.1** *A feasible allocation  $x^*$  is Pareto efficient if and only if  $x^*$  solves the following problem for all  $i = 1, 2, \dots, n$*

$$\begin{aligned} & \max_x u_i(x_i) \\ \text{s.t.} \quad & \sum x_k \leq \sum w_k \\ & u_k(x_k) \geq u_k(x_k^*) \text{ for } k \neq i \end{aligned}$$

Proof. Suppose  $x^*$  solves all maximizing problems but  $x^*$  is not Pareto efficient. This means that there is some allocation  $x'$  where one consumer is better off, and the others are not worse off. But, then  $x^*$  does not solve one of the maximizing problems, a contradiction. Conversely, suppose  $x^*$  is Pareto efficient, but it does not solve one of the problems. Instead, let  $x'$  solve that particular problem. Then  $x'$  makes one of the agents better off without hurting any of the other agents, which contradicts the assumption that  $x^*$  is Pareto efficient. The proof is completed. ■

If utility functions  $u_i(x_i)$  are differentiable, then we can define the Lagrangian function to get the optimal solution to the above problem:

$$L = u_i(x_i) + q(\hat{w} - \hat{x}) + \sum_{k \neq i} t_k [u_k(x_k) - u_k(x_k^*)]$$

The first order conditions are then given by

$$\frac{\partial L}{\partial x_i^l} = \frac{\partial u_i(x_i)}{\partial x_i^l} - q^l \leq 0 \quad \text{with equality if } x_i^l > 0 \quad l = 1, \dots, L, i = 1, \dots, n, \quad (8.6)$$

$$\frac{\partial L}{\partial x_k^l} = t_k \frac{\partial u_k(x_k)}{\partial x_k^l} - q^l \leq 0 \quad \text{with equality if } x_k^l > 0 \quad l = 1, \dots, L; k \neq i \quad (8.7)$$

By (8.6), when  $x^*$  is an interior solution, we have

$$\frac{\frac{\partial u_i(x_i)}{\partial x_i^l}}{\frac{\partial u_i(x_i)}{\partial x_i^h}} = \frac{q^l}{q^h} = MRS_{x_i^l, x_i^h} \quad (8.8)$$

By (8.7)

$$\frac{\frac{\partial u_k(x_k)}{\partial x_k^l}}{\frac{\partial u_k(x_k)}{\partial x_k^h}} = \frac{q^l}{q^h} \quad (8.9)$$

Thus, we have

$$MRS_{x_1^l, x_1^h} = \dots = MRS_{x_n^l, x_n^h} \quad l = 1, 2, \dots, L; h = 1, 2, \dots, L. \quad (8.10)$$

which are the necessary conditions for the interior solutions to be Pareto efficient, which means that the *MRS* of any two goods are all equal for all agents. They become sufficient conditions when utility functions  $u_i(x_i)$  are differentiable and quasi-concave.

## 8.4.2 Production Economies

For simplicity, we assume there is only one firm. Let  $T(y) \leq 0$  be the transformation frontier. Similarly, we can prove the following proposition.

**Proposition 8.4.2** *A feasible allocation  $(x^*, y^*)$  is Pareto efficient if and only if  $(x^*, y^*)$  solves the following problem for all  $i = 1, 2, \dots, n$*

$$\begin{aligned} & \max_x u_i(x_i) \\ \text{s.t.} \quad & \sum_{k \in N} x_k = \sum_{k \in N} w_k + y \\ & u_k(x_k) \geq u_k(x_k^*) \quad \text{for } k \neq i \\ & T(y) \leq 0. \end{aligned}$$

If utility functions  $u_i(x_i)$  are differentiable, then we can define the Lagrangian function to get the first order conditions:

$$L = u_i(x_i) - \lambda T(\hat{x} - \hat{w}) + \sum_{k \neq i} t_k [u_k(x_k) - u_k(x_k^*)]$$

FOC:

$$\frac{\partial L}{\partial x_i^l} = \frac{\partial u_i(x_i)}{\partial x_i^l} - \lambda^l \frac{\partial T(y)}{\partial x_i^l} \leq 0 \quad \text{with equality if } x_i^l > 0 \quad l = 1, \dots, L, i = 1, \dots, n, \quad (8.11)$$

$$\frac{\partial L}{\partial x_k^l} = t_k \frac{\partial u_k(x_k)}{\partial x_k^l} - \lambda^l \frac{\partial T(y)}{\partial x_k^l} \leq 0 \quad \text{with equality if } x_k^l > 0 \quad l = 1, 2, \dots, L; k \neq l \quad (8.12)$$

When  $x^*$  is an interior solution, we have by (8.11)

$$\frac{\frac{\partial u_i(x_i)}{\partial x_i^l}}{\frac{\partial u_i(x_i)}{\partial x_i^h}} = \frac{\frac{\partial T(y)}{\partial y^l}}{\frac{\partial T(y)}{\partial y^h}}, \quad (8.13)$$

and by (8.12)

$$\frac{\frac{\partial u_k(x_k)}{\partial x_k^l}}{\frac{\partial u_k(x_k)}{\partial x_k^h}} = \frac{\frac{\partial T(y)}{\partial y^l}}{\frac{\partial T(y)}{\partial y^h}}. \quad (8.14)$$

Thus, we have

$$MRS_{x_1^l, x_1^h} = \cdots = MRS_{x_n^l, x_n^h} = MRTS_{y^l, y^h} \quad l = 1, 2, \dots, L; h = 1, 2, \dots, L \quad (8.15)$$

which are the necessary condition for the interior solutions to be Pareto efficient, which means that the  $MRS$  of any two goods for all agents equals the  $MRTS$ . They become sufficient conditions when utility functions  $u_i(x_i)$  are differentiable and quasi-concave and the production functions are concave.

## 8.5 The Second Fundamental Theorem of Welfare Economics

Now we can assert a converse of the First Fundamental Theorem of Welfare Economics. The Second Fundamental Theorem of Welfare Economics gives conditions under which a Pareto optimum allocation can be “supported” by a competitive equilibrium if we allow some redistribution of endowments. It tells us, under some regularity assumptions, including essential condition of convexity of preferences and production sets, that any desired Pareto optimal allocation can be achieved as a market-based equilibrium with transfers.

We first define the competitive equilibrium with transfer payments (also called an equilibrium relative to a price system), that is, a competitive equilibrium is established after transferring some of initial endowments between agents.

**Definition 8.5.1 (Competitive Equilibrium with Transfer Payments)** For an economy,  $e = (e_1, \dots, e_n, \{Y_j\})$ ,  $(x, y, p) \in X \times Y \times R_+^L$  is a competitive equilibrium with transfer payment if

- (i)  $x_i \succsim_i x'_i$  for all  $x'_i \in \{x'_i \in X_i : px'_i \leq px_i\}$  for  $i = 1, \dots, n$ .
- (ii)  $py_j \geq py'_j$  for  $y'_j \in Y_j$ .
- (iii)  $\hat{x} \leq \hat{w} + \hat{y}$  (feasibility condition).

**Remark 8.5.1** An equilibrium with transfer payments is different from a competitive equilibrium with respect to a budget constrained utility maximization that uses a value



calculated from the initial endowment. An equilibrium with transfer payments is defined without reference to the distribution of initial endowment, given the total amount.

**Remark 8.5.2** If Condition (i) above is replaced by

$$(i') \ x'_i \succ_i x_i \text{ implies that } px'_i \geq px_i \text{ for } i = 1, \dots, n,$$

$(x, y, p) \in X \times Y \times R_+^L$  is called a price quasi-equilibrium with transfer payment. It may be remarked that, if  $\hat{w} > 0$ , preferences are continuous and locally non-satiated, then any price quasi-equilibrium with transfers in which  $x_i > 0$  is a price equilibrium with transfers. Furthermore, when preferences are strictly monotone, we must have  $p > 0$  in any price quasi-equilibrium with transfers. For the proofs, see Mas-Colell, Whinston, and Green (1995).

The following theorem which is called the Second Fundamental Theorem of Welfare Economics shows that every Pareto efficient allocation can be supported by a competitive equilibrium through a redistribution of endowments so that one does not need to seek any alternative economic institution to reach Pareto efficient allocations. This is one of the most important theorems in modern economics, and the theorem is also one of the theorems in microeconomic theory whose proof is complicated.

**Theorem 8.5.1 (The Second Fundamental Theorem of Welfare Economics)** *Suppose  $(x^*, y^*)$  with  $x^* > 0$  is Pareto optimal, suppose  $\succsim_i$  are continuous, convex and strictly monotonic, and suppose that  $Y_j$  are closed and convex. Then, there is a price vector  $p \geq 0$  such that  $(x^*, y^*, p)$  is a competitive equilibrium with transfer payments, i.e.,*

$$(1) \text{ if } x'_i \succ_i x_i^*, \text{ then } px'_i > px_i^* \text{ for } i = 1, \dots, n.$$

$$(2) \text{ } py_j^* \geq py'_j \text{ for all } y'_j \in Y_j \text{ and } j = 1, \dots, J.$$

Proof: Let

$$P(x_i^*) = \{x_i \in X_i : x_i \succsim_i x_i^*\} \tag{8.16}$$

be the strict upper contour set and let

$$P(x^*) = \sum_{i=1}^n P(x_i^*). \tag{8.17}$$

By the convexity of  $\succsim_i$ , we know that  $P(x_i^*)$  and thus  $P(x^*)$  are convex.

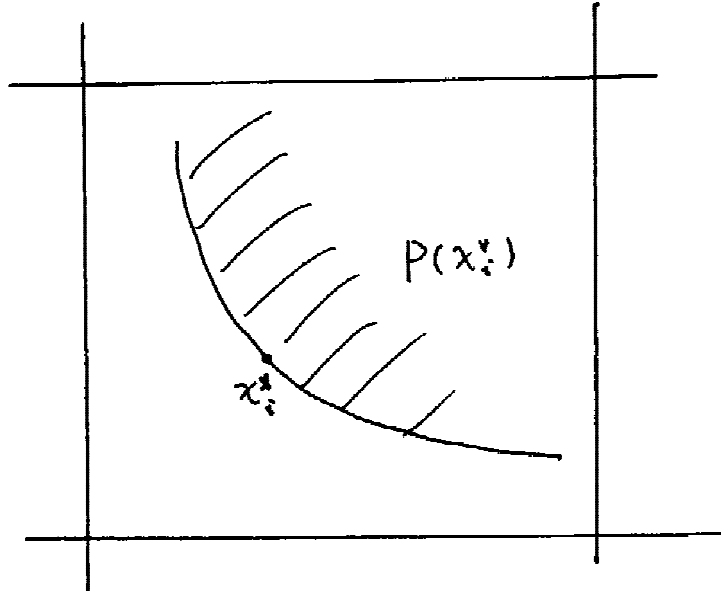


Figure 8.7:  $P(x_i^*)$  is the set of all points strictly above the indifference curve through  $x_i^*$ .

Let  $W = \{\hat{w}\} + \sum_{j=1}^J Y_j$  which is closed and convex. Then  $W \cap P(x^*) = \emptyset$  by Pareto optimality of  $(x^*, y^*)$ , and thus, by the Separating Hyperplane Theorem in Chapter 1, there is a  $p \neq 0$  such that

$$p\hat{z} \geq p\hat{\sigma} \text{ for all } \hat{z} \in P(x^*) \text{ and } \hat{\sigma} \in W \quad (8.18)$$

Now we show that  $p$  is a competitive equilibrium price vector by the following four steps.

1.  $p \geq 0$

To see this, let  $e^l = (0, \dots, 1, 0, \dots, 0)$  with the  $l$ -th component one and other places zero. Let

$$\hat{z} = \hat{\sigma} + e^l \quad \text{for some } \hat{\sigma} \in W.$$

Then  $\hat{z} = \hat{\sigma} + e^l \in P(x^*)$  by strict monotonicity and redistribution of  $e^l$ . Thus, we have by (8.18).

$$p(\hat{\sigma} + e^l) \geq p\hat{\sigma} \quad (8.19)$$

and thus

$$pe^l \geq 0 \quad (8.20)$$

which means

$$p^l \geq 0 \quad \text{for } l = 1, 2, \dots, L. \quad (8.21)$$

2.  $py_j^* \geq py_j$  for all  $y_j \in Y_j$ ,  $j = 1, \dots, J$ .

Since  $\hat{x}^* = \hat{y}^* + \hat{w}$  by noting  $(x^*, y^*)$  is a Pareto efficient allocation and preference orderings are strictly monotonic, we have  $p\hat{x}^* = p(\hat{w} + \hat{y}^*)$ . Thus, by  $p\hat{z} \geq p(\hat{w} + \hat{y})$  in (8.18) and  $p\hat{w} = p\hat{x}^* - p\hat{y}^*$ , we have

$$p(\hat{z} - \hat{x}^*) \geq p(\hat{y} - \hat{y}^*) \quad \forall \hat{y} \in \hat{Y}.$$

Letting  $\hat{z} \rightarrow x^*$ , we have

$$p(\hat{y} - \hat{y}^*) \leq 0 \quad \forall \hat{y} \in \hat{Y}.$$

Letting  $y_k = y_k^*$  for  $k \neq j$ , we have from the above equation,

$$py_j^* \geq py_j \quad \forall y_j \in Y_j.$$

3. If  $x_i \succ_i x_i^*$ , then

$$px_i \geq px_i^*. \quad (8.22)$$

To see this, let

$$\begin{aligned} x'_i &= (1 - \theta)x_i \quad 0 < \theta < 1 \\ x'_k &= x_k^* + \frac{\theta}{n-1}x_i \quad \text{for } k \neq i \end{aligned}$$

Then, by the continuity of  $\succ_i$  and the strict monotonicity of  $\succ_k$ , we have  $x'_i \succ_i x_i^*$  for all  $i \in N$ , and thus

$$x' \in P(x^*) \quad (8.23)$$

if  $\theta$  is sufficiently small. By (8.18), we have

$$p(x'_i + \sum_{k \neq i} x'_k) = p[(1 - \theta)x_i + \sum_{k \neq i} (x_k^* + \frac{\theta}{n-1}x_i)] \geq p \sum_{k=1}^n x_k^* \quad (8.24)$$

and thus we have

$$px_i \geq px_i^* \quad (8.25)$$

4. If  $x_i \succ_i x_i^*$ , we must have  $px_i > px_i^*$ . To show this, suppose by way of contradiction, that

$$px_i = px_i^* \quad (8.26)$$

Since  $x_i \succ_i x_i^*$ , then  $\lambda x_i \succ_i x_i^*$  for  $\lambda$  sufficiently close to one by the continuity of preferences for  $0 < \lambda < 1$ . By step 3, we know  $\lambda px_i \geq px_i^* = px_i$  so that  $\lambda \geq 1$  by  $px_i = px_i^* > 0$ , which contradicts the fact that  $\lambda < 1$ .

**Remark 8.5.3** When  $\hat{w} > 0$ , the strict monotonicity condition of preferences can be weakened to locally non-satiated. Also, Mas-Colell, Whinston, and Green (1995) use price quasi-equilibrium concept to present a version of Second Fundamental Theorem of Welfare Economics under much weaker conditions, which only requires the convexities of preferences and production possibilities and local non-satiation of preferences. Thus, when  $\hat{w} > 0$ , we can obtain the Second Theorem of Welfare Economics under the local non-satiation, continuity, and convexity of preferences. For the discussion, see Mas-Colell, Whinston, and Green (1995).

For exchange economies, the competitive equilibrium with transfers is the same as a regular competitive equilibrium when  $w_i = x_i^*$ . As a corollary, we have

**Corollary 8.5.1** *Suppose  $x^* > 0$  is Pareto optimal, suppose  $\succsim_i$  are continuous, convex and strictly monotonic. Then,  $x^*$  is a competitive equilibrium for the initial endowment  $w_i = x_i^*$ .*

**Remark 8.5.4** If  $\succsim_i$  can be represented by a concave and differentiable utility function, then the proof of the Second Fundamental Theorem of Welfare Economics can be much simpler. A sufficient condition for concavity is that the Hessian matrix is negative definite. Also note that monotonic transformation does not change preferences so that we may be able to transform a quasi-concave utility function to a concave utility function as follows, for example

$$\begin{aligned} u(x, y) &= xy \text{ which is quasi-concave} \\ \Leftrightarrow u^{\frac{1}{2}}(x, y) &= x^{\frac{1}{2}}y^{\frac{1}{2}} \end{aligned}$$

which is concave after monotonic transformation.

### Differentiation Version of the Second Fundamental Theorem of Welfare Economics for Exchange Economies

Proof: If  $x^* > 0$  is Pareto Optimal, then we have

$$Du_i(x_i) = q \quad i = 1, 2, \dots, n. \quad (8.27)$$

We want to show  $q$  is a competitive equilibrium price vector. To do so, we only need to show that each consumer maximizes his utility s.t.  $B(q) = \{x_i \in X_i : qx_i \leq qx_i^*\}$ . Indeed, by concavity of  $u_i$

$$\begin{aligned} u_i(x_i) &\leq u(x_i^*) + Du_i(x_i^*)(x_i - x_i^*) \\ &= u(x_i^*) + q(x_i - x_i^*)/t_i \\ &\leq u(x_i^*) \end{aligned}$$

The reason the inequality holds for a concave function is because that, from Figure 3.8, we have

$$\frac{u(x) - u(x^*)}{x - x^*} \leq u'(x^*). \quad (8.28)$$

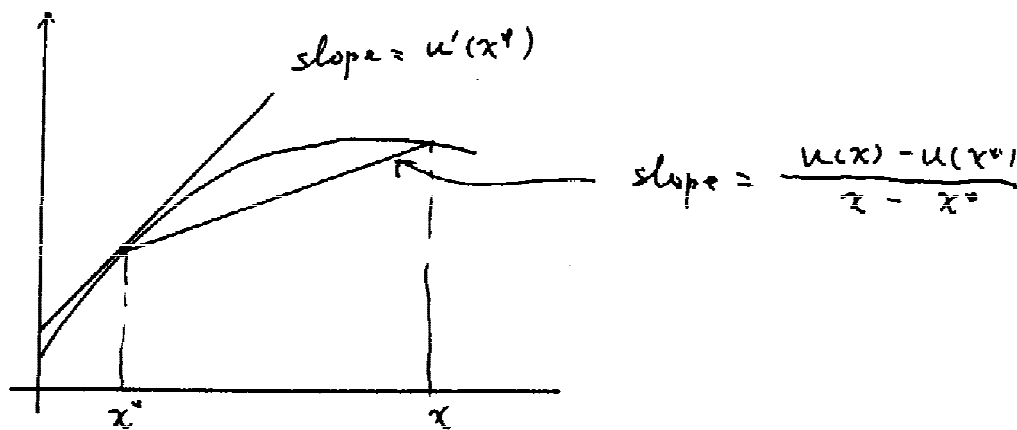


Figure 8.8: Concave function

Thus, we have  $u_i(x_i) \leq u(x_i^*) + Du_i(x_i^*)(x_i - x_i^*)$ .

## 8.6 Non-Convex Production Technologies and Marginal Cost Pricing

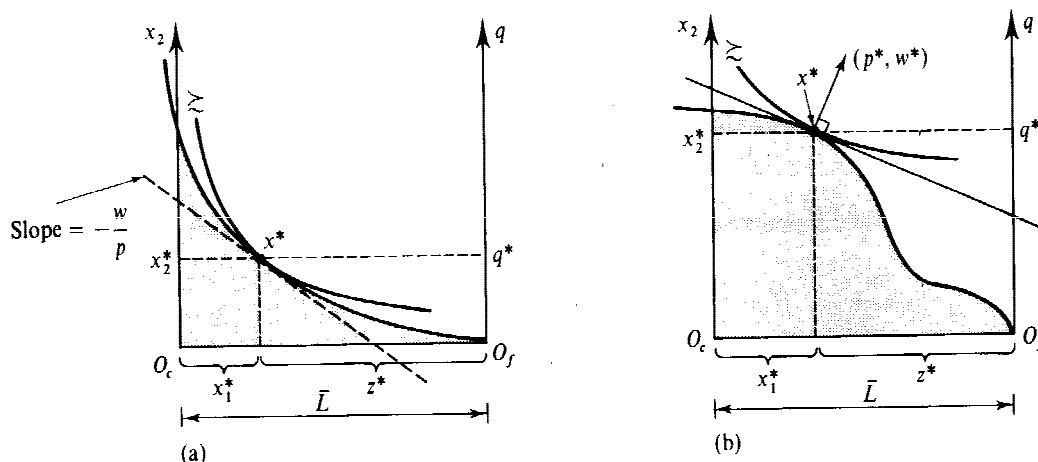


Figure 8.9: Figure (a) shows failure of the second welfare theorem with a non-convex technology. Figure (b) shows the first welfare theorem applies even with a non-convex technology.

The indispensability of convexity for the second welfare theorem can be observed in Figure 8.9(a). There, the allocation  $x^*$  maximizes the welfare of the consumer, but for the only value of relative prices that could support  $x^*$  as a utility-maximizing bundle, the firm does not maximize profits even locally (i.e., at the relative prices  $w/p$ , there are productions arbitrarily close to  $x^*$  yielding higher profits). In contrast, the first welfare theorem remains applicable even in the presence of non-convexities. As Figure 8.9(b) suggests, any Walrasian equilibrium maximizes the well-being of the consumer in the feasible production set.

The second welfare theorem runs into difficulties in the presence of non-convex production sets (here we do not question the assumption of convexity on the consumption side). In the first place, large non-convexities caused by the presence of fixed costs or extensive increasing returns lead to a world of a small number of large firms, making the assumption of price taking less plausible. Yet, even if price taking can somehow be relied on, it may still be impossible to support a given Pareto optimal allocation. Examples are provided by Figures 8.9(a) and 8.10. In Figure 8.10, at the only relative prices that could

support the production  $y^*$  locally, the firm sustains a loss and would rather avoid it by shutting down. In Figure 8.9(a), on the other hand, not even local profit maximization can be guaranteed.

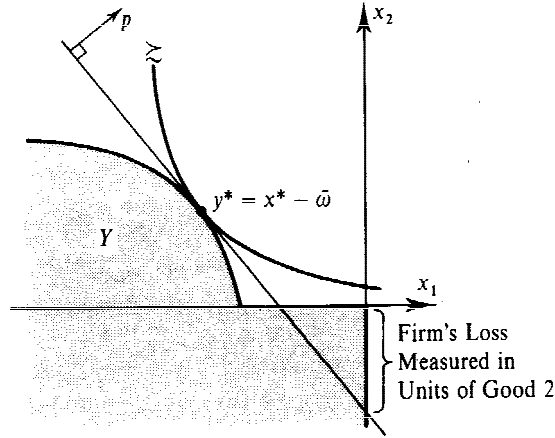


Figure 8.10: The firm incurs a loss at the prices that locally support be Pareto optimal allocation.

Although non-convexities may prevent us from supporting the Pareto optimal production allocation as a profit-maximizing choice, under the differentiability assumptions on pricing rules of firms, we can use the first-order necessary conditions derived there to formulate a weaker result that parallels the second welfare theorem.

**Theorem 8.6.1 (The Second Welfare Theorem for Non-Convex Production Sets)**

Suppose that firm's  $j$  transformation function  $T_j(y)$  are smooth, and all consumers have continuous, monotonic, and convex preferences. If  $(x^*, y^*)$  is Pareto optimal, then there exists a price vector  $p$  and wealth levels  $(I_1, \dots, I_n)$  with  $\sum_i I_i = p \cdot \sum_i w_i + \sum_j p \cdot y_j^*$  such that:

(i) For any firm  $j$ , we have

$$p = \gamma_j \nabla T_j(y^*) \text{ for some } \gamma_j > 0.$$

(ii) For any  $i$ ,  $x_i^*$  is maximal for  $\succsim_i$  in the budget set

$$\{x_i \in X_i' : p \cdot x_i \leq I_i\}.$$

(iii)  $\sum_i x_i^* = \sum_i w_i + \sum_j y_j^*$ .

The type of equilibrium represented by conditions (i) to (iii) of the above theorem is called a *marginal cost price equilibrium with transfers*. The motivation for this terminology comes from the one-output, one-input case.

Condition (i) also implies that

$$\frac{p^l}{p^k} = \frac{\frac{\partial T_1^l(y)}{\partial y^l}}{\frac{\partial T_1(y)}{\partial y^k}} = \dots = \frac{\frac{\partial T_J^l(y)}{\partial y^l}}{\frac{\partial T_J(y)}{\partial y^k}}, \quad (8.29)$$

and thus marginal rates of technic substitutions, *MRTS*, of two goods are equal for all firms. As we have noted, although condition (i) is necessary, but not sufficient, for profit maximization, and in fact, it does not imply that the  $(y_1^*, \dots, y_J^*)$  are profit-maximizing production plans for price-taking firms. The condition says only that small changes in production plans have no first-order effect on profit. But small changes may still have positive second-order effects (as in Figure 8.9(a), where at a marginal cost price equilibrium the firm actually chooses the production that *minimizes* profits among the efficient productions) and, at any rate, large changes may increase profits (as in Figure 8.10). Thus, to achieve allocation  $(x^*, y^*)$  may require that a regulatory agency prevent the managers of non-convex firms from attempting to maximize profits at the given prices  $p$ . Recently, Tian (2010) presented such an incentive compatible mechanism that implements marginal cost price equilibrium allocations with transfers in Nash equilibrium. Tian (2009) also consider implementation of other equilibrium allocations when firms pricing rule is given by such as average-cost pricing, loss-free pricing, etc.

It should be noted that the converse result to the above theorem, which would assert that every marginal cost price equilibrium is Pareto optimal, is not true. In Figure 8.11, for example, we show a one-consumer economy with a non-convex production set. In the figure,  $x^*$  is a marginal cost price equilibrium with transfers for the price system  $p = (1, 1)$ . Yet, allocation  $x'$  yields the consumer a higher utility. Informally, this occurs because marginal cost pricing neglects second-order conditions and it may therefore happen that, as at allocation  $x^*$ , the second-order conditions for the social utility maximization problem are not. A marginal cost pricing equilibrium need not be Pareto optimal allocation. As a result, satisfaction of the first-order marginal optimality conditions (which in the case of Figure 8.11 amounts simply to the tangency of the indifference curve and the production surface) does not ensure that the allocation is Pareto optimal.



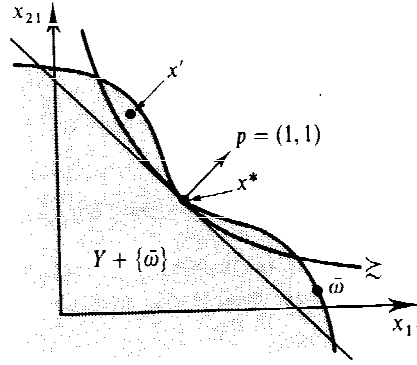


Figure 8.11: The firm incurs a loss at the prices that locally support be Pareto optimal allocation.

See Quinzii (1992) for extensive background and discussion on the material presented in this section.

## 8.7 Pareto Optimality and Social Welfare Maximization

Pareto efficiency is only concerned with efficiency of allocations and has nothing to say about distribution of welfare. Even if we agree with Pareto optimality, we still do not know which one we should be at. One way to solve the problem is to assume the existence of a social welfare function.

Define a social welfare function  $W : X \rightarrow \Re$  by  $W(u_1(x_1), \dots, u_n(x_n))$  where we assume that  $W(\cdot)$  is monotone increasing.

**Example 8.7.1 (Utilitarian Social Welfare Function)** The utilitarian social welfare function is given by

$$W(u_1, \dots, u_n) = \sum_{i=1}^n a_i u_i(x_i)$$

with  $\sum a_i = 1, a_i \geq 0$ . Under a utilitarian rule, social states are ranked according to the linear sum of utilities. The utilitarian form is by far the most common and widely applied social welfare function in economics.

**Example 8.7.2 (Rawlsian Social Welfare Function)** The Rawlsian social welfare function is defined by

$$W(\cdot) = \min\{u_1(x_1), u_2(x_2), \dots, u_n(x_n)\}.$$

So the utility function is not strictly monotonic increasing. The Rawlsian form gives priority to the interests of the worst off members, and it is used in the ethical system proposed by Rawls (1971).

### 8.7.1 Social Welfare Maximization for Exchange Economies

We suppose that a society should operate at a point that maximizes social welfare; that is, we should choose an allocation  $x^*$  such that  $x^*$  solves

$$\max W(u_1(x_1), \dots, u_n(x_n))$$

subject to

$$\sum_{i=1}^n x_i \leq \sum_{i=1}^n w_i.$$

How do the allocations that maximize this welfare function compare to Pareto efficient allocations? The following is a trivial consequence if the strict monotonicity assumption is imposed.

**Proposition 8.7.1** *Under strict monotonicity of the social welfare function,  $W(\cdot)$  if  $x^*$  maximizes a social welfare function, then  $x^*$  must be Pareto Optimal.*

Proof: If  $x^*$  is not Pareto Optimal, then there is another feasible allocation  $x'$  such that  $u_i(x'_i) \geq u_i(x_i)$  for all  $i$  and  $u_k(x'_k) > u_k(x_k)$  for some  $k$ . Then, by strict monotonicity of  $W(\cdot)$ , we have  $W(u_1(x'_1), \dots, u_n(x'_n)) > W(u_1(x_1), \dots, u_n(x_n))$  and thus it does not maximize the social welfare function. ■

Thus, every social welfare maximum is Pareto efficient. Is the converse necessarily true? By the Second Fundamental Theorem of Welfare Economics, we know that every Pareto efficient allocation is a competitive equilibrium allocation by redistributing endowments. This gives us a further implication of competitive prices, which are the multipliers for the welfare maximization. Thus, the competitive prices really measure the (marginal) social value of a good. Now we state the following proposition that shows every Pareto efficient allocation is a social welfare maximum for the social welfare function with a suitable weighted sum of utilities.

**Proposition 8.7.2** *Let  $x^* > 0$  be a Pareto optimal allocation. Suppose  $u_i$  is concave, differentiable and strictly monotonic. Then, there exists some choice of weights  $a_i^*$  such that  $x^*$  maximizes the welfare functions*

$$W(u_1, \dots, u_n) = \sum_{i=1}^n a_i u_i(x_i) \quad (8.30)$$

Furthermore,  $a_i^* = \frac{1}{\lambda_i}$  with  $\lambda_i = \frac{\partial V_i(p, I_i)}{\partial I_i}$

where  $V_i(\cdot)$  is the indirect utility function of consumer  $i$ .

Proof: Since  $x^*$  is Pareto optimal, it is a competitive equilibrium allocation with  $w_i = x_i^*$  by the second theorem of welfare economies. So we have

$$D_i u_i(x_i^*) = \lambda p \quad (8.31)$$

by the first order condition, where  $p$  is a competitive equilibrium price vector.

Now for the welfare maximization problem

$$\begin{aligned} & \max \sum_{i=1}^n a_i u_i(x_i) \\ & s.t. \sum x_i \leq \sum x_i^* \end{aligned}$$

since  $u_i$  is concave,  $x^*$  solves the problem if the first order condition

$$a_i \frac{\partial u_i(x_i)}{\partial x_i} = q \quad i = 1, \dots, n \quad (8.32)$$

is satisfied for some  $q$ . Thus, if we let  $p = q$ , then  $a_i^* = \frac{1}{\lambda_i}$ . We know  $x^*$  also maximizes the welfare function  $\sum_{i=1}^n a_i^* u_i(x_i')$ . ■

Thus, the price is the Lagrangian multiples of the welfare maximization, and this measures the marginal social value of goods.

## 8.7.2 Welfare Maximization in Production Economy

Define a choice set by the transformation function

$$T(\hat{x}) = 0 \quad \text{with} \quad \hat{x} = \sum_{i=1}^n x_i. \quad (8.33)$$

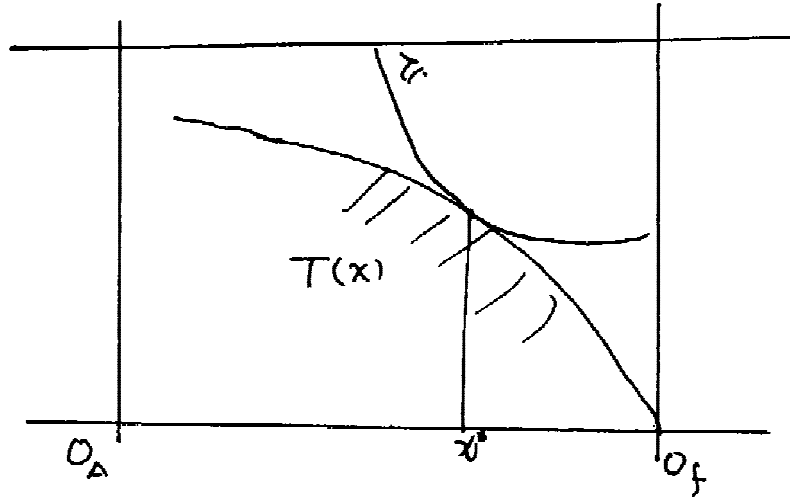


Figure 8.12: Welfare maximization

The social welfare maximization problem for the production economy is

$$\max W(u_1(x_1), u_2(x_2), \dots, u_n(u_n)) \quad (8.34)$$

subject to

$$T(\hat{x}) = 0.$$

Define the Lagrangian function

$$L = W(u_1(x_1), \dots, u_n(u_n)) - \lambda T(\hat{x}). \quad (8.35)$$

The first order condition is then given by

$$W'(\cdot) \frac{\partial u_i(x_i)}{\partial x_i^l} - \lambda \frac{\partial T(\hat{x})}{\partial x_i^l} \leq 0 \quad \text{with equality if } x_i^l > 0, \quad (8.36)$$

and thus when  $x$  is an interior point, we have

$$\frac{\frac{\partial u_i(x_i)}{\partial x_i^l}}{\frac{\partial u_i(x_i)}{\partial x_i^k}} = \frac{\frac{\partial T(\hat{x})}{\partial x_i^l}}{\frac{\partial T(\hat{x})}{\partial x_i^k}} \quad (8.37)$$

That is,

$$MRS_{x_i^l, x_i^k} = MRTS_{x_i^l, x_i^k}. \quad (8.38)$$

The conditions characterizing welfare maximization require that the marginal rate of substitution between each pair of commodities must be equal to the marginal rate of transformation between the two commodities for all agents.

## 8.8 Political Overtones

1. By the First Fundamental Theorem of Welfare Economics, implication is that what the government should do is to secure the competitive environment in an economy and give people full economic freedom. So, as long as the market system works well, there should be no subsidizing, no price floor, no price ceiling, stop a rent control, no regulations, lift the tax and the import-export barriers.

2. Even if we want to reach a preferred Pareto optimal outcome which may be different from a competitive equilibrium from a given endowment, you might do so by adjusting the initial endowment but not disturbing prices, imposing taxes or regulations. That is, if a derived Pareto optimal is not “fair”, all the government has to do is to make a lump-sum transfer payments to the poor first, keeping the competitive environments intact. We can adjust initial endowments to obtain a desired competitive equilibrium by the Second Fundamental Theorem of Welfare Economics.

3. Of course, when we reach the above conclusions, you should note that there are conditions on the results. In many cases, we have market failures, in the sense that either a competitive equilibrium does not exist or a competitive equilibrium may not be Pareto optimal so that the First or Second Fundamental Theorem of Welfare Economics cannot be applied.

The conditions for the existence of a competitive equilibrium are: (i) convexity (diversification of consumption and no IRS), (ii) monotonicity (self-interest), and (iii) continuity, (iv) divisibility, (v) perfect competition, (vi) complete information, etc. If these conditions are not satisfied, we may not obtain the existence of a competitive equilibrium. The conditions for the First Fundamental Theorem of Welfare Economics are: (i) local non-satiation (unlimited desirability), (ii) divisibility, (iii) no externalities, (iv) perfect competition, (v) complete information etc. If these conditions are not satisfied, we may not guarantee that every competitive equilibrium allocation is Pareto efficient. The conditions for the Second Fundamental Theorem of Welfare Economics are: (i) the convexity of preferences and production sets, (ii) monotonicity (self-interest), and (iii) continuity, (iv) divisibility, (v) perfect competition, (vi) complete information, etc. If these conditions are not satisfied, we may not guarantee every Pareto efficient allocation can be supported by a competitive equilibrium with transfers.

Thus, as a general notice, before making an economic statement, one should pay attention to the assumptions which are implicit and/or explicit involved. As for a conclusion from the general equilibrium theory, one should notice conditions such as divisibility, no externalities, no increasing returns to scale, perfect competition, complete information. If these assumptions are relaxed, a competitive equilibrium may not exist or may not be Pareto efficient, or a Pareto efficient allocation may not be supported by a competitive equilibrium with transfer payments. Only in this case of a market failure, we may adopt another economic institution. We will discuss the market failure and how to solve the market failure problem in Part IV and Part V.

## Reference

- Arrow, K and G. Debreu, "Existence of Equilibrium for a Competitive Economy," *Econometrica*, 1954, 22.
- Arrow, K. J. and Hahn, F. H. (1971), *General Competitive Analysis*, (San Francisco, Holden Day).
- Debreu, G. (1959), *Theory of Value*, (Wiley, New York).
- Hildenbrand, W., A. P. Kirman, *Equilibrium Analysis: Variations on the Thems by Edgeworth and Walras*, North-Holland: New York, 1988.
- Jehle, G. A., and P. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, 1998, Chapter 7.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapters 6, 10.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995, Chapters 10, 16.
- Quinzii, M. *Increasing Returns and Efficiency*, Oxford University Press, New York, 1992.
- Rawls, J., *A Theory of Justice*, Cambridge: Harvard University Press, 1971.
- Takayama, A. *Mathematical Economics*, the second edition, Cambridge: Cambridge University Press, 1985, Chapters 1-3.

- Tian, G., "On the Constrained Walrasian and Lindahl Correspondences," *Economics Letters*, 26 (1988), pp. 299-303.
- Tian, G., "Implementation in Economies with Non-Convex Production Technologies Unknown to the Designer," *Games and Economic Behavior*, 66 (2009), 526-545.
- Tian, G., "Implementation of Marginal Cost Pricing Equilibrium Allocations with Transfers in Economies with Increasing Returns to Scale," *Review of Economic Design* , 14 (2010), 163-184.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 17-18.

# Chapter 9

## Economic Core, Fair Allocations, and Social Choice Theory

### 9.1 Introduction

In this chapter we briefly discuss some topics in the framework of general equilibrium theory, namely economic core, fair allocations, and social choice theory. The theory of core is important because it gives an insight into how a competitive equilibrium is achieved as a result of individual strategic behavior instead of results of an auctioneer and the Walrasian tâtonnement mechanism. It shows the necessity of adopting a market institution as long as individuals behave self-interestedly.

We have also seen that Pareto optimality may be too weak a criterion to be meaningful. It does not address any question about income distribution and equity of allocations. Fairness is a notion to overcome this difficulty. This is one way to restrict a set of Pareto optimum.

In a slightly different framework, suppose that a society is deciding the social priority among finite alternatives. Alternatives may be different from Pareto optimal allocations. Let us think of a social “rule” to construct the social ordering (social welfare function) from many individual orderings of different alternatives. The question is: Is it possible to construct a rule satisfying several desirable properties? Both “fairness” and “social welfare function” address a question of social justice.



## 9.2 The Core of Exchange Economies

The use of a competitive (market) system is just one way to allocate resources. What if we use some other social institution? Would we still end up with an allocation that was “close” to a competitive equilibrium allocation? The answer will be that, if we allow agents to form coalitions, the resulting allocation can only be a competitive equilibrium allocation when the economy becomes large. Such an allocation is called a core allocation and was originally considered by Edgeworth (1881).

The core is a concept in which every individual and every group agree to accept an allocation instead of moving away from the social coalition.

There is some reason to think that the core is a meaningful political concept. If a group of people find themselves able, using their own resources to achieve a better life, it is not unreasonable to suppose that they will try to enforce this threat against the rest of community. They may find themselves frustrated if the rest of the community resorts to violence or force to prevent them from withdrawing.

The theory of the core is distinguished by its parsimony. Its conceptual apparatus does not appeal to any specific trading mechanism nor does it assume any particular institutional setup. Informally, the notion of competition that the theory explores is one in which traders are well informed of the economic characteristics of other traders, and in which the members of any group of traders can bind themselves to any mutually advantageous agreement.

For simplicity, we consider exchange economies. We say two agents are of the same type if they have the same preferences and endowments.

The  $r$ -replication of the original economy: There are  $r$  times as many agents of each type in the original economy.

A *coalition* is a group of agents, and thus it is a subset of  $n$  agents.

**Definition 9.2.1 (Blocking Coalition)** A group of agents  $S$  (a coalition) is said to block (improve upon) a given allocation  $x$  if there is a Pareto improvement with their own resources, that is, if there is some allocation  $x'$  such that

- (1) it is feasible for  $S$ , ie.,  $\sum_{i \in S} x'_i \leq \sum_{i \in S} w_i$ ,
- (2)  $x'_i \succ_i x_i$  for all  $i \in S$  and  $x'_k \succ_k x_k$  for some  $k \in S$ .

**Definition 9.2.2 (Core)** A feasible allocation  $x$  is said to have the core property if it cannot be improved upon for any coalition. The core of an economy is the set of all allocations that have the core property.

**Remark 9.2.1** Every allocation in the core is Pareto optimal (coalition by whole people).

**Definition 9.2.3 (Individual Rationality)** An allocation  $x$  is individually rational if  $x_i \succsim_i w_i$  for all  $i = 1, 2, \dots, n$ .

The individual rationality condition is also called the participation condition which means that a person will not participate the economic activity if he is worse off than at the initial endowment.

**Remark 9.2.2** Every allocation in the core must be individually rational.

**Remark 9.2.3** When  $n = 2$  and preference relations are weakly monotonic, an allocation is in the core if and only if it is Pareto optimal and individually rational.

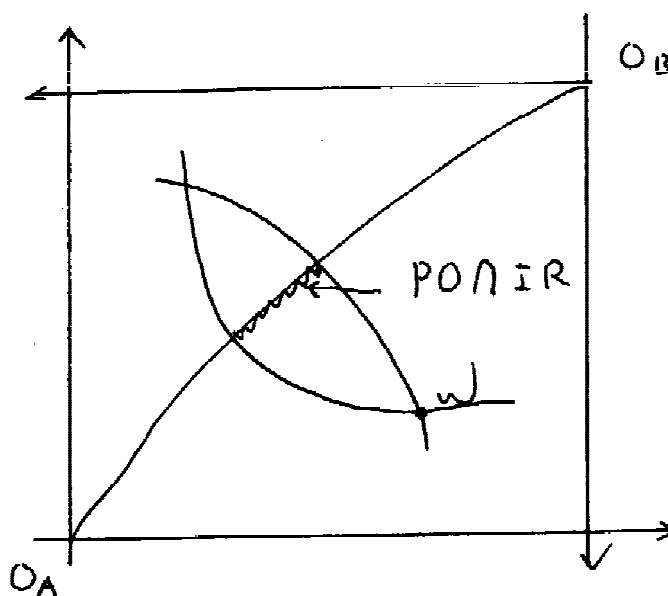


Figure 9.1: The set of allocations in the core is simply given by the set of Pareto efficient and individually rational allocations when  $n = 2$ .

**Remark 9.2.4** Even though a Pareto optimal allocation is independent of individual endowments, an allocation in the core depends on individual endowments.

What is the relationship between core allocations and competitive equilibrium allocations?

**Theorem 9.2.1** *Under local non-satiation, if  $(x, p)$  is a competitive equilibrium, then  $x$  has the core property.*

Proof: Suppose  $x$  is not an allocation in the core. Then there is a coalition  $S$  and a feasible allocation  $x'$  such that

$$\sum_{i \in S} x'_i \leq \sum_{i \in S} w_i \quad (9.1)$$

and  $x'_i \succ_i x_i$  for all  $i \in S$ ,  $x'_k \succ_k x_k$  for some  $k \in S$ . Then, by local non-satiation, we have

$$\begin{aligned} px'_i &\geq px_i \text{ for all } i \in S \text{ and} \\ px'_k &> px_k \text{ for some } k \end{aligned}$$

Therefore, we have

$$\sum_{i \in S} px'_i > \sum_{i \in S} px_i = \sum_{i \in S} pw_i \quad (9.2)$$

a contradiction. Therefore, the competitive equilibrium must be an allocation in the core.

We refer to the allocations in which consumers of the same type get the same consumption bundles as *equal-treatment allocations*. It can be shown that any allocation in the core must be an equal-treatment allocation.

**Proposition 9.2.1 (Equal Treatment in the Core)** *Suppose agents' preferences are strictly convex. Then if  $\mathbf{x}$  is an allocation in the  $r$ -core of a given economy, then any two agents of the same type must receive the same bundle.*

Proof: Let  $x$  be an allocation in the core and index the  $2r$  agents using subscripts  $A1, \dots, Ar$  and  $B1, \dots, Br$ . If all agents of the same type do not get the same allocation, there will be one agent of each type who is most poorly treated. We will call these two agents the “type-A underdog” and the “type-B underdog.” If there are ties, select any of the tied agents.

Let  $\bar{x}_A = \frac{1}{r} \sum_{j=1}^r x_{A_j}$  and  $\bar{x}_B = \frac{1}{r} \sum_{j=1}^r x_{B_j}$  be the average bundle of the type-A and type-B agents. Since the allocation  $x$  is feasible, we have

$$\frac{1}{r} \sum_{j=1}^r x_{A_j} + \frac{1}{r} \sum_{j=1}^r x_{B_j} = \frac{1}{r} \sum_{j=1}^r \omega_{A_j} + \frac{1}{r} \sum_{j=1}^r \omega_{B_j} = \frac{1}{r} r\omega_A + \frac{1}{r} r\omega_B.$$

It follows that

$$\bar{x}_A + \bar{x}_B = \omega_A + \omega_B,$$

so that  $(\bar{x}_A, \bar{x}_B)$  is feasible for the coalition consisting of the two underdogs. We are assuming that at least for one type, say type A, two of the type-A agents receive different bundles. Hence, the A underdog will strictly prefer  $\bar{x}_A$  to his present allocation by strict convexity of preferences (since it is a weighted average of bundles that are at least as good as  $x_A$ ), and the B underdog will think  $\bar{x}_B$  is at least as good as his present bundle, thus forming a coalition that can improve upon the allocation. The proof is completed.

Since any allocation in the core must award agents of the same type with the same bundle, we can examine the cores of replicated two-agent economies by use of the Edgeworth box diagram. Instead of a point  $x$  in the core representing how much A gets and how much B gets, we think of  $x$  as telling us how much each agent of type A gets and how much each agent of type B gets. The above proposition tells us that all points in the  $r$ -core can be represented in this manner.

The following theorem is a converse of Theorem 9.2.1 and shows that any allocation that is not a market equilibrium allocation must eventually not be in the  $r$ -core of the economy. This means that core allocations in large economies look just like Walrasian equilibria.

**Theorem 9.2.2 (Shrinking Core Theorem)** *Suppose  $\succsim_i$  are strictly convex and continuous. Suppose  $x^*$  is a unique competitive equilibrium allocation. Then, if  $y$  is not a competitive equilibrium, there is some replication  $V$  such that  $y$  is not in the  $V$ -core.*

Proof: We prove the theorem with the aid of Figure 9.2. From the figure, one can see that  $y$  is not a competitive equilibrium. We want to show that there is a coalition such that the point  $y$  can be improved upon for  $V$ -replication. Since  $y$  is not a competitive equilibrium, the line segment through  $y$  and  $w$  must cut at least one agent's, say agent A's, indifference curve through  $y$ . Then, by strict convexity and continuity of  $\succsim_i$ , there are integers  $V$  and  $T$  with  $0 < T < V$  such that

$$g_A \equiv \frac{T}{V}w_A + \left(1 - \frac{T}{V}\right)y_A \succ_A y_A.$$

We can do so since any real number can be approached by a rational number that consists of the ratio of two suitable integers  $\frac{T}{V}$ .

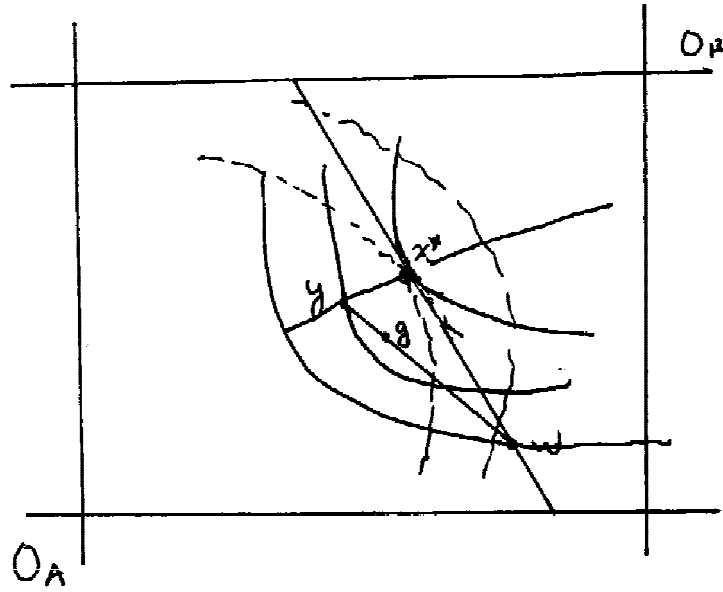


Figure 9.2: The shrinking core. A point like  $y$  will eventually not be in the core.

Now, form a coalition consisting of  $V$  consumers of type A and  $V - T$  consumers of type B and consider the allocation  $x = (g_A, \dots, g_A, y_B, \dots, y_B)$  (in which there are  $V$  Type A and  $V - T$  Type B). We want to show  $x$  is feasible for this coalition.

$$\begin{aligned}
 Vg_A + (V - T)y_B &= V \left[ \frac{T}{V}w_A + \left( 1 - \frac{T}{V}y_A \right) \right] + (V - T)y_B \\
 &= Tw_A + (V - T)y_A + (V - T)y_B \\
 &= Tw_A + (V - T)(y_A + y_B) \\
 &= Tw_A + (V - T)(w_A + w_B) \\
 &= Vw_A + (V - T)w_B
 \end{aligned}$$

by noting  $y_A + y_B = w_A + w_B$ . Thus,  $x$  is feasible in the coalition and  $g_A \succ_A y_A$  for all agents in type A and  $y_B \sim_B y_B$  for all agents in type B which means  $y$  is not in the  $V$ -core for the  $V$ -replication of the economy. The proof is completed.

**Remark 9.2.5** The shrinking core theorem then shows that the only allocations that are in the core of a large economy are market equilibrium allocations, and thus Walrasian

equilibria are robust: even very weak equilibrium concepts, like that of core, tend to yield allocations that are close to Walrasian equilibria for larger economies. Thus, this theorem shows the essential importance of competition and fully economic freedom.

**Remark 9.2.6** Many of the restrictive assumptions in this proposition can be relaxed such as strict monotonicity, convexity, uniqueness of competitive equilibrium, and two types of agents.

From the above discussion, we have the following limit theorem.

**Theorem 9.2.3 (Limit Theorem on the Core)** *Under the strict convexity and continuity, the core of a replicated two person economy shrinks when the number of agents for each type increases, and the core coincides with the competitive equilibrium allocation if the number of agents goes to infinity.*

This result means that any allocation which is not a competitive equilibrium allocation is not in the core for some  $r$ -replication.

## 9.3 Fairness of Allocation

Pareto efficiency gives a criterion of how the goods are allocated efficiently, but it may be too weak a criterion to be meaningful. It does not address any questions about income distribution, and does not give any “equity” implication. Fairness is a notion that may overcome this difficulty. This is one way to restrict the whole set of Pareto efficient outcomes to a small set of Pareto efficient outcomes that satisfy the other properties.

What is the equitable allocation?

How can we define the notion of equitable allocation?

**Definition 9.3.1 (Envy)** An agent  $i$  is said to envy agent  $k$  if agent  $i$  prefers agent  $k$ 's consumption. i.e.,  $x_k \succ_i x_i$ .

**Definition 9.3.2** An allocation  $x$  is equitable if no one envies anyone else, i.e., for each  $i \in N$ ,  $x_i \succsim_i x_k$  for all  $k \in N$ .

**Definition 9.3.3 (Fairness)** An allocation  $x$  is said to be fair if it is both Pareto optimal and equitable.



The following figure shows that  $x$  is Pareto efficient, but not equitable.

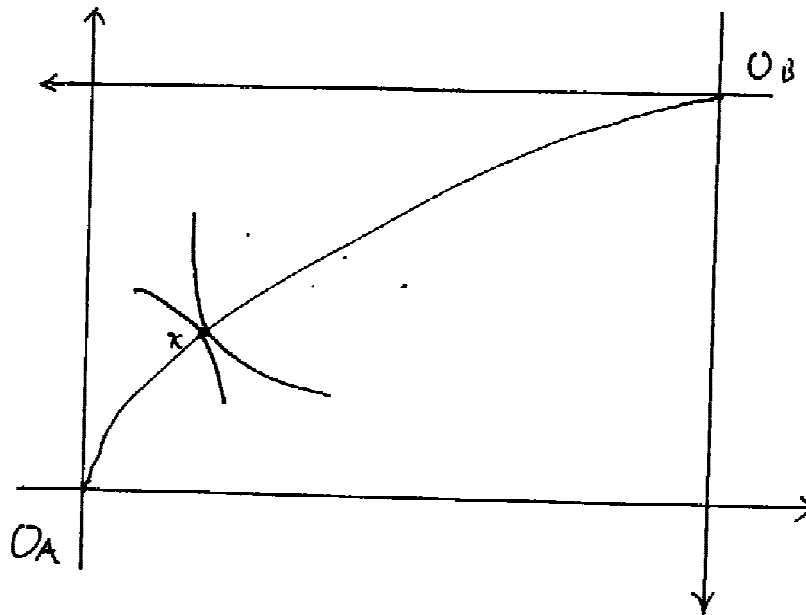


Figure 9.4:  $x$  is Pareto efficient, but not equitable.

The figure below shows that  $x$  is equitable, but it is not Pareto efficient.

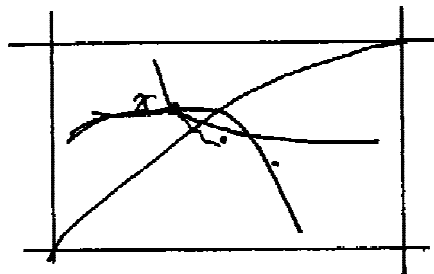


Figure 9.5:  $x$  is equitable, but not Pareto efficient.

How to test a fair allocation?

### Graphical Procedures for Testing Fairness:

Let us restrict an economy to a two-person economy. An easy way for agent  $A$  to compare his own allocation  $x_A$  with agent  $B$ 's allocation  $x_B$  in the Edgeworth Box is to find a point symmetric of  $x_A$  against the center of the Box. That is, draw a line from  $x_A$  to the center of the box and extrapolate it to the other side by the same length to find



$x'_A$ , and then make the comparison. If the indifference curve through  $x_A$  cuts “below”  $x'_A$ , then  $A$  envies  $B$ . Then we have the following way to test whether an allocation is a fair allocation:

- Step 1: Is it Pareto optimality? If the answer is “yes”, go to step 2; if no, stop.
- Step 2: Construct a reflection point  $(x_B, x_A)$ . (Note that  $\frac{x_A + x_B}{2}$  is the center of the Edgeworth box.)
- Step 3: Compare  $x_B$  with  $x_A$  for person  $A$  to see if  $x_B \succ_A x_A$  and compare  $x_A$  with  $x_B$  for person  $B$  to see if  $x_A \succ_B x_B$ . If the answer is “no” for both persons, it is a fair allocation.

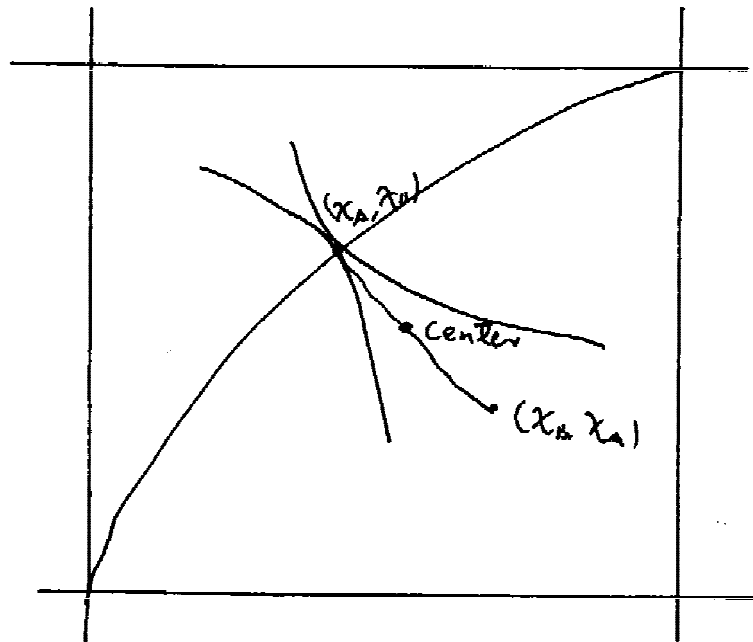


Figure 9.6: How to test a fair allocation.

We have given some desirable properties of “fair” allocations. A question is whether it exists at all. The following theorem provides one sufficient condition to guarantee the existence of fairness.

**Theorem 9.3.1** *Let  $(x^*, p^*)$  be a competitive equilibrium. Under local non-satiation, if all individuals’ income is the same, i.e.,  $p^*w_1 = p^*w_2 = \dots = p^*w_n$ , then  $x^*$  is a strictly fair allocation.*

Proof: By local non-satiation,  $x^*$  is Pareto optimal by the First Fundamental Theorem of Welfare Economics. We only need to show  $x^*$  is strictly equitable. Suppose not. There is  $i$  and a coalition  $S$  with  $i \notin S$  such that

$$\bar{x}_S^* \equiv \frac{1}{|S|} \sum_{k \in S} x_k^* \succ_i x_i^* \quad (9.3)$$

Then, we have  $p^* \bar{x}_S^* > p^* x_i^* = p^* w_i$ . But this contradicts the fact that

$$p^* \bar{x}_S^* = \frac{1}{|S|} \sum_{k \in S} p^* x_k^* = \frac{1}{|S|} \sum_{k \in S} p^* w_k = p^* w_i \quad (9.4)$$

by noting that  $p^* w_1 = p^* w_2 = \dots = p^* w_n$ . Therefore, it must be a strictly fair allocation.

**Definition 9.3.6** An allocation  $x \in \mathfrak{R}_+^{nL}$  is an equal income Walrasian allocation if there exists a price vector such that

- (1)  $px_i \leq p\bar{w}$ , where  $\bar{w} = \frac{1}{n} \sum_{k=1}^n w_k$  : average endowment.
- (2)  $x'_i \succ_i x_i$  implies  $px'_i > p\bar{w}$ .
- (3)  $\sum x_i \leq \sum w_i$ .

Notice that every equal Walrasian allocation  $x$  is a competitive equilibrium allocation with  $w_i = \bar{w}$  for all  $i$ .

**Corollary 9.3.1** *Under local non-satiation, every equal income Walrasian allocation is strictly fair allocation.*

**Remark 9.3.5** An “equal” division of resource itself does not give “fairness,” but trading from “equal” position will result in “fair” allocation. This “divide-and-choose” recommendation implies that if the center of the box is chosen as the initial endowment point, the competitive equilibrium allocation is fair. A political implication of this remark is straightforward. Consumption of equal bundle is not Pareto optimum, if preferences are different. However, an equal division of endowment plus competitive markets result in fair allocations.

**Remark 9.3.6** A competitive equilibrium from an equitable (but not “equal” division) endowment is not necessarily fair.

**Remark 9.3.7** Fair allocations are defined without reference to initial endowments. Since we are dealing with optimality concepts, initial endowments can be redistributed among agents in a society.

In general, there is no relationship between an allocation in the core and a fair allocation. However, when the social endowments are divided equally among two persons, we have the following theorem.

**Theorem 9.3.2** *In a two-person exchange economy, if  $\succsim_i$  are convex, and if the total endowments are equally divided among individuals, then the set of allocations in the core is a subset of (strictly) fair allocation.*

Proof: We know that an allocation in the core  $x$  is Pareto optimal. We only need to show that  $x$  is equitable. Since  $x$  is a core allocation, then  $x$  is individually rational (everyone prefers initial endowments). If  $x$  is not equitable, there is some agent  $i$ , say, agent  $A$ , such that

$$\begin{aligned} x_B \succ_A x_A \succsim_A w_A &= \frac{1}{2}(w_A + w_B) \\ &= \frac{1}{2}(x_A + x_B) \end{aligned}$$

by noting that  $w_A = w_B$  and  $x$  is feasible. Thus,  $x_A \succsim_A \frac{1}{2}(x_A + x_B)$ . But, on the other hand, since  $\succsim_A$  is convex,  $x_B \succ_A x_A$  implies that  $\frac{1}{2}(x_A + x_B) \succ_A x_A$ , a contradiction.

## 9.4 Social Choice Theory

### 9.4.1 Introduction

In this section, we present a very brief summary and introduction of social choice theory. We analyze the extent to which individual preferences can be aggregated into social preferences, or more directly into social decisions, in a “satisfactory” manner (in a manner compatible with the fulfilment of a variety of desirable conditions).

As was shown in the discussion of “fairness,” it is difficult to come up with a criterion (or a constitution) that determines that society’s choice. Social choice theory aims at constructing such a rule which could be allied with not only Pareto efficient allocations,

but also any alternative that a society faces. We will give some fundamental results of social choice theory: Arrow Impossibility Theorem, which states there does not exist any non-dictatorial social welfare function satisfying a number of “reasonable” assumptions. Gibbard-Satterthwaite theorem states that no social choice mechanism exists which is non-dictatorial and can never be advantageously manipulated by some agents.

### 9.4.2 Basic Settings

To simplify the exposition, we will return to the notation used in pre-orders of preferences. Thus  $aP_ib$  means that  $i$  strictly prefers  $a$  to  $b$ . We will also assume that the individual preferences are all strict, in other words, that there is never a situation of indifference. For all  $i$ ,  $a$  and  $b$ , one gets preference  $aP_ib$  or  $bP_ia$  (this hypothesis is hardly restrictive if  $A$  is finite).

$N = \{1, 2, \dots, n\}$  : the set of individuals;

$X = \{x_1, x_2, \dots, x_m\} : (m \geq 3)$ : the set of alternatives (outcomes);

$P_i = (\text{or } \succ_i)$  : strict preference orderings of agent  $i$ ;

$\mathcal{P}_i$  the class of allowed individual orderings;

$P = (P_1, P_2, \dots, P_n)$  : a preference ordering profile;

$\mathcal{P}$ : the set of all profiles of individuals orderings;

$S(X)$  : the class of allowed social orderings.

Arrow’s social welfare function:

$$F : \mathcal{P} \rightarrow S(X) \tag{9.5}$$

which is a mapping from individual ordering profiles to social orderings.

Gibbard-Satterthwaite’s social choice function (SCF) is a mapping from individual preference orderings to the alternatives

$$f : \mathcal{P} \rightarrow X \tag{9.6}$$

Note that even though individuals’ preference orderings are transitive, a social preference ordering may not be transitive. To see this, consider the following example.

**Example 9.4.1 (The Condorcet Paradox)** Suppose a social choice is determined by the majority voting rule. Does this determine a well-defined social welfare function? The answer is in general no by the well-known Condorcet paradox. Consider a society with three agents and three alternatives:  $x, y, z$ . Suppose each person's preference is given by

$$x \succ_1 y \succ_1 z \quad (\text{by person 1})$$

$$y \succ_2 z \succ_2 x \quad (\text{by person 2})$$

$$z \succ_3 x \succ_3 y \quad (\text{by person 3})$$

By the majority rule,

$$\text{For } x \text{ and } y, \quad xFy \text{ (by social preference)}$$

$$\text{For } y \text{ and } z, \quad yFz \text{ (by social preference)}$$

$$\text{For } x \text{ and } z, \quad zFx \text{ (by social preference)}$$

Then, pairwise majority voting tells us that  $x$  must be socially preferred to  $y$ ,  $y$  must be socially preferred to  $z$ , and  $z$  must be socially preferred to  $x$ . This cyclic pattern means that social preference is not transitive.

The number of preference profiles can increase very rapidly with increase of number of alternatives.

**Example 9.4.2**  $X = \{x, y, z\}$ ,  $n = 3$

$$x \succ y \succ z$$

$$x \succ z \succ y$$

$$y \succ x \succ z$$

$$y \succ z \succ x$$

$$z \succ x \succ y$$

$$z \succ y \succ x$$

Thus, there are six possible individual orderings, i.e.,  $|\mathcal{P}_i| = 6$ , and therefore there are  $|\mathcal{P}_1| \times |\mathcal{P}_2| \times |\mathcal{P}_3| = 6^3 = 216$  possible combinations of 3-individual preference orderings on three alternatives. The social welfare function is a mapping from each of these 216 entries (cases) to one particular social ordering (among six possible social orderings of three alternatives). The social choice function is a mapping from each of these 216 cases to one particular choice (among three alternatives). A question we will investigate is what kinds of desirable conditions should be imposed on these social welfare or choice functions.

You may think of a hypothetical case that you are sending a letter of listing your preference orderings,  $P_i$ , on announced national projects (alternatives), such as reducing deficits, expanding the medical program, reducing society security program, increasing national defence budget, to your Congressional representatives. The Congress convenes with a huge stake of letters  $P$  and try to come up with national priorities  $F(P)$ . You want the Congress to make a certain rule (the Constitution) to form national priorities out of individual preference orderings. This is a question addressed in social choice theory.

### 9.4.3 Arrow's Impossibility Theorem

**Definition 9.4.1** Unrestricted Domain (UD): A class of allowed individual orderings  $\mathcal{P}_i$  consists of all possible orderings defined on  $X$ .

**Definition 9.4.2** Pareto Principle (P): if for  $x, y \in X$ ,  $x P_i y$  for all  $i \in N$ , then  $x F(P) y$  (social preferences).

**Definition 9.4.3** Independence of Irrelevant Alternatives (IIA): For any two alternatives  $x, y \in X$  and any two preference profiles  $P, P' \in \mathcal{P}^n$ ,  $x F(P) y$  and  $\{i : x P_i y\} = \{i : x P'_i y\}$  for all  $i$  implies that  $x F(P') y$ . That is, the social preference between any two alternatives depends only on the profile of individual preferences over the same alternatives.

**Remark 9.4.1** IIA means that the ranking between  $x$  and  $y$  for any agent is equivalent in terms of  $P$  and  $P'$  implies that the social ranking between  $x$  and  $y$  by  $F(P)$  and  $F(P')$  is the same. In other words, if two different preference profiles that are the same on  $x$  and  $y$ , the social order must be also the same on  $x$  and  $y$ .

**Remark 9.4.2** By IIA, any change in preference ordering other than the ordering of  $x$  and  $y$  should not affect social ordering between  $x$  and  $y$ .

**Example 9.4.3** Suppose  $x P_i y P_i z$  and  $x P'_i z P'_i y$ .

By IIA, if  $x F(P) y$ , then  $x F(P') y$ .

**Definition 9.4.4 (Dictator)** There is some agent  $i \in N$  such that  $F(P) = P_i$  for all  $P \in [\mathcal{P}^n$ , and agent  $i$  is called a dictator.

**Theorem 9.4.1 (Arrow's Impossibility Theorem)** *Any social welfare function that satisfies  $m \geq 3$ , UD, P, IIA conditions is dictatorial.*

Proof: The idea for the following proof dates back to Vickrey (1960); it comprises three lemmas.

**Lemma 9.4.1 (Neutrality)** *Given a division  $N = M \cup I$  and  $(a, b, x, y) \in A^4$  such that*

$$\cdot \forall i \in M, x P_i y \text{ and } a P'_i b$$

$$\cdot \forall i \in I, y P_i x \text{ and } b P'_i a$$

*then*

$$x \hat{P} y \Leftrightarrow a \hat{P}' b \quad (9.7)$$

$$x \hat{I} y \Leftrightarrow a \hat{I}' b \quad (9.8)$$

Proof: The interpretation of lemma 9.4.1 is simple: If  $x$  and  $y$  are ordered by each individual in  $P$  as  $a$  and  $b$  in  $P'$ , then this must also be true of social preference; that is,  $x$  and  $y$  must be ordered by  $\hat{P}$  as  $a$  and  $b$  by  $\hat{P}'$ . If this weren't the case, then the procedure of aggregation would treat the pairs  $(a, b)$  and  $(x, y)$  in a non-neutral manner, hence the name of the lemma.

Let us first prove (9.7). Suppose that  $x \hat{P}' y$  and that  $(a, b, x, y)$  are distinct two by two. Let preferences  $P''$  be such that

$$\cdot \forall i \in M, a P''_i x P''_i y P''_i b$$

$$\cdot \forall i \in I, y P''_i b P''_i a P''_i x$$

(such preferences exist because the domain is universal). By IIA, since  $x$  and  $y$  are ordered individually in  $P''$  as in  $P$ , we have  $x \hat{P}'' y$ . By the Pareto principle,  $a \hat{P}'' x$  and  $y \hat{P}'' b$ . By transitivity, we get  $a \hat{P}'' b$ . Finally, in reapplying IIA, we find that  $a \hat{P}' b$ . The cases where  $a$  or  $b$  coincide with  $x$  or  $y$  are treated similarly.

Part (9.8) is obtained directly. If  $x \hat{I} y$  but, for example  $a \hat{P}' b$ , we find the contradiction  $x \hat{P} y$  by (9.7) in reversing the roles of  $(a, b)$  and  $(x, y)$ .

Before stating Lemma 9.4.2, two terms must be defined. We say that a set of agents  $M$  is almost decisive on  $(x, y)$  if for every  $\hat{P}$ ,

$$(\forall i \in M, xP_iy \text{ and } \forall i \notin M, yP_ix) \Rightarrow x\hat{P}y$$

We say that  $M$  is decisive on  $(x, y)$  if for every  $\hat{P}$ ,

$$(\forall i \in M, xP_iy) \Rightarrow x\hat{P}y$$

**Lemma 9.4.2** *If  $M$  is almost decisive on  $(x, y)$ , then it is decisive on  $(x, y)$ .*

Proof: Suppose that  $\forall i \in M, xP_iy$ . By IIA, only individual preferences count on  $(x, y)$ ; the others can be changed. Assume therefore, with no loss of generality, that  $z$  exists such that  $xP_izP_iy$  if  $i \in M$  and  $z$  is preferred to  $x$  and  $y$  by the other agents. Neutrality imposes that  $x\hat{P}z$  because the individual preferences are oriented as on  $(x, y)$  and  $M$  is almost decisive on  $(x, y)$ . Finally, the Pareto principle implies that  $z\hat{P}y$  and transitivity implies the conclusion that  $x\hat{P}y$ .

Note that neutrality implies that if  $M$  is decisive on  $(x, y)$ , it is decisive on every other pair. We will therefore just say that  $M$  is decisive.

**Lemma 9.4.3** *If  $M$  is decisive and contains at least two agents, then a strict subset of  $M$  exists that is decisive.*

Proof: Divide  $M = M_1 \cup M_2$ , and choose a  $\hat{P}$  such that

- on  $M_1, xP_iyP_iz$
- on  $M_2, yP_izP_ix$
- outside  $M, zP_ixP_iy$

As  $M$  is decisive, we have  $y\hat{P}z$ .

One of two things results:

- Either  $y\hat{P}x$ , and (since  $z$  doesn't count by IIA)  $M_2$  is almost decisive on  $(x, y)$ , and therefore decisive by lemma 9.4.2.
- Or  $x\hat{P}y$ , and by transitivity  $x\hat{P}z$ ; then ( $y$  doesn't count by IIA)  $M_1$  is almost decisive on  $(x, y)$ , and therefore is decisive by lemma refsoc2.



The proof is concluded by noting that by the Pareto principle,  $I$  as a whole is decisive. In apply Lemma 9.4.3, the results is an individual  $i$  that is decisive, and this is therefore the sought-after dictator. The attentive reader will have noted that lemma 9.4.1 remains valid when there are only two alternatives, contrary to Lemmas 9.4.2 and 9.4.3. The proof of the theorem is completed.

The impact of the Arrow possibility theorem has been quite substantial. Obviously, Arrow's impossibility result is a disappointment. The most pessimistic reaction to it is to conclude that there is just no acceptable way to aggregate individual preferences, and hence no theoretical basis for treating welfare issues. A more moderate reaction, however, is to examine each of the assumptions of the theorem to see which might be given up. Conditions imposed on social welfare functions may be too restrictive. Indeed, when some conditions are relaxed, then the results could be positive, e.g., UD is usually relaxed.

#### 9.4.4 Some Positive Result: Restricted Domain

When some of the assumptions imposed in Arrow's impossibility theorem is removed, the result may be positive. For instance, if alternatives have certain characteristics which could be placed in a spectrum, preferences may show some patterns and may not exhaust all possibilities orderings on  $X$ , thus violating (UD). In the following, we consider the case of restricted domain. A famous example is a class of "single-peaked" preferences. We show that under the assumption of single-peaked preferences, non-dictatorial aggregation is possible.

**Definition 9.4.5** A binary relation  $\geq$  on  $X$  is a *linear order* if  $\geq$  is reflexive ( $x \geq x$ ), transitive ( $x \geq y \geq z$  implies  $x \geq z$ ), and total (for distinct  $x, y \in X$ , either  $x \geq y$  or  $y \geq x$ , but not both).

Example:  $X = \Re$  and  $x \geq y$ .

**Definition 9.4.6**  $\succsim_i$  is said to be *single-peaked* with respect to the linear order  $\geq$  on  $X$ , if there is an alternative  $x \in X$  such that  $\succsim_i$  is increasing with respect to  $\geq$  on the lower contour set  $L(x) = \{y \in X : y \leq x\}$  and decreasing with respect to  $\geq$  on the upper contour set  $U(x) = \{y \in X : y \geq x\}$ . That is,

(1)  $x \geq z > y$  implies  $z \succ_i y$

(2)  $y > z \geq x$  implies  $z \succ_i y$

In words, there is an alternative that represents a peak of satisfaction and, moreover, satisfaction increases as we approach this peak so that, in particular, there cannot be any other peak of satisfaction.

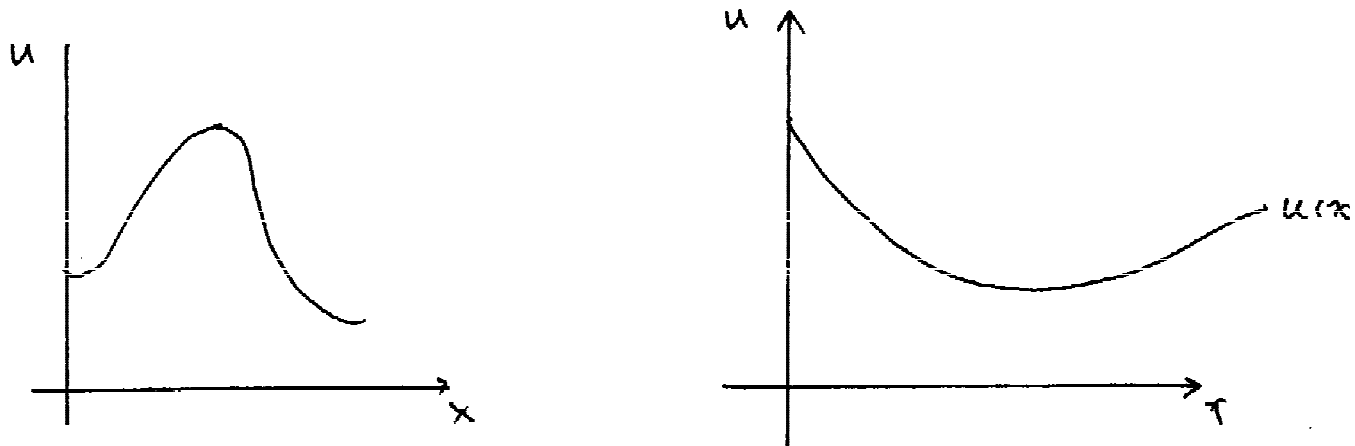


Figure 9.7:  $u$  in the left figure is single-peaked,  $u$  in the right figure is not single-peaked.

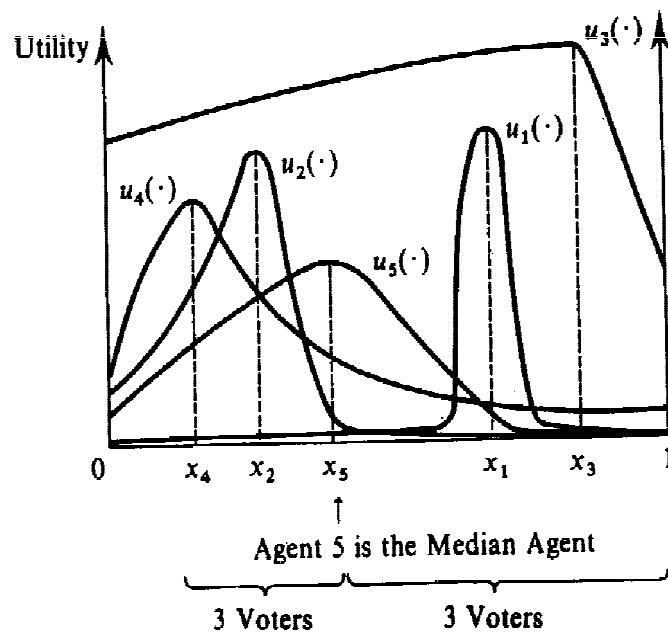


Figure 9.8: Five Agents who have single-peaked preferences.

Given a profile of preference  $(\succsim_1, \dots, \succsim_n)$ , let  $x_i$  be the maximal alternative for  $\succsim_i$  (we will say that  $x_i$  is “individual  $i$ ’s peak”).

**Definition 9.4.7** Agent  $h \in N$  is a median agent for the profile  $(\succsim_1, \dots, \succsim_n)$  if  $\#\{i \in N : x_i \geq x_h\} \geq \frac{n}{2}$  and  $\#\{i \in N : x_h \geq x_i\} \geq \frac{n}{2}$ .

**Proposition 9.4.1** Suppose that  $\geq$  is a linear order on  $X$ ,  $\succsim_i$  is single-peaked. Let  $h \in N$  be a median agent, then the majority rule  $\tilde{F}(\succsim)$  is aggregatable:

$$x_h \tilde{F}(\succsim) y \quad \forall y \in X.$$

That is, the peak  $x_h$  of the median agent is socially optimal (cannot be defeated by any other alternative) by majority voting. Any alternative having this property is called a Condorect winner. Therefore, a Condorect winner exists whenever the preferences of all agents are single-peaked with respect to the same linear order.

Proof. Take any  $y \in X$  and suppose that  $x_h > y$  (the argument is the same for  $y > x_h$ ). We need to show that

$$\#\{i \in N : x_h \succ_i y\} \geq \#\{i \in N : y \succ_i x_h\}.$$

Consider the set of agents  $S \subset N$  that have peaks larger than or equal to  $x_h$ , that is,  $S = \{i \in N : x_i \geq x_h\}$ . Then  $x_i \geq x_h > y$  for every  $i \in S$ . Hence, by single-peakness of  $\succsim_i$  with respect to  $\geq$ , we get  $x_h \succ_i y$  for every  $i \in S$  and thus  $\#S \leq \#\{i \in N : x_h \succ_i y\}$ . Hence,  $\{i \in N : y \succ_i x_h\} \subset (N \setminus S)$ , and then  $\#\{i \in N : y \succ_i x_h\} \leq \#(N \setminus S)$ . On the other hand, because agent  $h$  is a median agent, we have that  $\#S \geq n/2$  and therefore  $\#\{i \in N : y \succ_i x_h\} \leq \#(N \setminus S) \leq n/2 \leq \#S \leq \#\{i \in N : x_h \succ_i y\}$ .

### 9.4.5 Gibbard-Satterthwaite Impossibility Theorem

The task we set ourselves to accomplish in the previous sections was how to aggregate profiles of individuals preference relations into a rational social preference order. Presumably, this social preference order is then used to make decisions. In this section we focus directly on social decision and pose the aggregation question as one analyzing how profiles of individual preferences turn into social decisions.

The main result we obtain again yields a dictatorship conclusion. The result amounts in a sense, to a translation of the Arrow's impossibility theorem into the language of choice functions. It provides a link towards the incentive-based analysis in the mechanism design we will discuss in Part V.

**Definition 9.4.8** A social choice function (SCF) is manipulable at  $P \in \mathcal{P}^n$  if there exists  $P'_i \in \mathcal{P}$  such that

$$f(P_{-i}, P'_i) P_i f(P_{-i}, P_i) \quad (9.9)$$

where  $P_{-i} = (P_1, \dots, P_{i-1}, P_{i+1}, \dots, P_n)$ .

**Definition 9.4.9** A SCF is strongly individually incentive compatible (SIIC) if there exists no preference ordering profile at which it is manipulable. In other words, the truth telling is a dominant strategy equilibrium:

$$f(P'_{-i}, P_i) P_i f(P'_{-i}, P'_i) \quad \text{for all } P' \in \mathcal{P}^n \quad (9.10)$$

**Definition 9.4.10** A SCF is dictatorial if there exists an agent whose optimal choice is the social optimal.

**Theorem 9.4.2 (Gibbard-Satterthwaite Theorem)** *If  $X$  has at least 3 alternatives, a SCF which is SIIC and UD is dictatorial.*

Proof. There are several the ways to prove the Gibbard-Satterthwaite's Impossibility Theorem. The proof adopted here is due to Schmeidler-Sonnenschein (1978). It has the advantage of bringing to light a close relationship to Arrow's Impossibility Theorem.

We want to prove that if such a mechanism is manipulable, then  $f$  is dictatorial. The demonstration comprises two lemmas; it consists of starting from an implementable SCF  $f$  and of constructing from it an SWF  $F$  that verifies Arrow's conditions. One concludes from this that  $F$  is dictatorial, which implies that  $f$  also is.

**Lemma 9.4.4** *Suppose that  $f(P) = a_1$  and  $f(P'_i, P_{-i}) = a_2$ , where  $a_2 \neq a_1$ . Then*

1.  *$f$  is manipulable by  $i$  in  $(P'_i, P_{-i})$  if  $a_1 P'_i a_2$*
2.  *$f$  is manipulable by  $i$  in  $P$  if  $a_2 P_i a_1$*

*Proof* In both cases it is sufficient to write the definition of manipulability.

We will need the notation  $P_i^j$ , which, for a profile  $P$  and given agents  $i < j$ , will represent the vector  $(P_i, \dots, P_j)$ .

LEMMA 2 Let  $B$  be a subset of the image of  $f$  and  $P$  a profile such that

- $\forall a_1 \in B, \forall a_2 \notin B, \forall i = 1, \dots, n, a_1 P_i a_2$

Then  $f(P) \in B$ .

*Proof* This can be shown by contradiction. Let  $a_2 = f(P)$ , and suppose that  $a_2 \notin B$ . Let  $P'$  be a profile such that  $f(P') = a_1 \in B$  (such a profile does exist, since  $B$  is included in the image of  $f$  and given the universal domain hypothesis). Now construct a sequence  $(a_3^i)_{i=0, \dots, n}$  by

- $a_3^0 = a_2 \notin B$   
for  $i = 1, \dots, n-1$ ,  $a_3^i = f(P_1^i, P_{i+1}^n)$   
 $a_3^n = a_1 \in B$

Let  $j$  be the first integer such that  $a_3^j \in B$ . We then get

- $f(P_1^j, P_{j+1}^n) = a_3^j \in B$   
 $f(P_1^{j-1}, P_j^n) = a_3^{j-1} \notin B$

and by the hypothesis of the lemma,  $a_3^j P_j a_3^{j-1}$ . Lemma 1 then implies that  $f$  is manipulable.

Now construct an SWF  $F$ . Let  $P$  be any profile and  $a_1, a_2$  two choices in  $A$ . Define a new profile (using UD)  $\tilde{P}$  such that for each  $i$ ,

- $\tilde{P}_i$ , coincides with  $P_i$ , on  $\{a_1, a_2\}$
- $\tilde{P}_i$ , coincides with  $P_i$ , on  $A - \{a_1, a_2\}$
- $\{a_1, a_2\}$  is placed at the top of the preferences  $\tilde{P}_i$

(Strictly speaking,  $\tilde{P}$  of course depends on  $a_1$  and  $a_2$ , and the notation should reflect this.)

Lemma 2 implies that  $f(\tilde{P}) \in \{a_1, a_2\}$  (taking  $B = \{a_1, a_2\}$  and replacing  $P$  by  $\tilde{P}$  in the statement of the lemma).  $F$  can therefore be defined by

- $a_1 F(P) a_2 \Leftrightarrow f(\tilde{P}) = a_1$

Now we can verify Arrow's conditions: There are surely at least three choices.

- $F$  is, by construction, of universal domain.
- $F$  satisfies the Pareto principle: if for every  $i$ ,  $a_1 P_i a_2$ , then  $a_1$  is at the top of all preferences  $\tilde{P}_i$ . By taking  $B = \{a_1, a_2\}$  in the statement of lemma 2, we indeed get  $f(\tilde{P}) = a_1$ .
- $F$  satisfies *IIA*: if this were not the case, there would exist  $P, P', a_1$  and  $a_2$  such that

for every  $i$ ,  $a_1 P_i a_2 \Leftrightarrow a_1 P'_i a_2$

- $a_1 F(P) a_2$  and  $a_2 F(P) a_1$

Now define a sequence  $(a_3^i)_{i=0, \dots, n}$  by

- $a_3^n = a_1$   
for  $i = 1, \dots, n-1$ ,  $a_3^i = f(\tilde{P}_1^n, \tilde{P}_{i+1}^n)$   
 $a_3^0 = a_2$

Lemma 2 implies that  $a_3^i \in \{a_1, a_2\}$  for every  $i$ . Therefore let  $j$  be the first integer such that  $a_3^j = a_2$ . This gives  $f(\tilde{P}_1^j, \tilde{P}_{j+1}^n) = a_2$  and  $f(\tilde{P}_1^{j-1}, \tilde{P}_j^n) = a_1$ . Now one of two things can result:

- $a_1 P_j a_2$

This implies  $a_1 P'_j a_2$  and therefore  $a_1 \tilde{P}'_j a_2$ , so lemma 1 implies that  $f$  is manipulable.

- $a_2 P_j a_1$

This implies  $a_2\tilde{P}_ja_1$ , so lemma 1 again implies that  $f$  is manipulable.

But there is contradiction in both results: for every  $P$ ,  $F(P)$  is clearly a complete and asymmetrical binary relation. What remains for us is to verify that it is transitive.

Take the opposite case so that we have a cycle on a triplet  $\{a_1, a_2, a_3\}$ . For every  $i$ , let  $P'_i$ , which coincides with  $P_i$ , on  $\{a_1, a_2, a_3\}$  and on  $A - \{a_1, a_2, a_3\}$  be such that  $\{a_1, a_2, a_3\}$  is at the top of  $P'_i$  (using UD). Lemma 2 implies that  $f(P') \in \{a_1, a_2, a_3\}$ ; without any loss of generality, we can assume that  $f(P') = a_1$ . Since  $F(P)$  has a cycle on  $\{a_1, a_2, a_3\}$ , we necessarily get  $a_2F(P)a_1$  or  $a_3F(P)a_1$ . Here again, without loss of generality, we can assume that  $a_3F(P)a_1$ . Now modify  $P'$  in  $P''$  by making  $a_2$  move into third place in each individual preference ( $P''$  is admissible by UD). Note that  $a_3P_ia_1$  if and only if  $a_3P''_ia_1$ ; in applying *IIA* (which we have just shown is satisfied), we get  $a_3F(P'')a_1$ , which again implies  $a_3 = f(P'')$ .

At the risk of seeming redundant, we now define a sequence  $(a_4^i)_{i=0,\dots,n}$  by

- $a_4^0 = a_1$
- for  $i = 1, \dots, n-1$ ,  $a_4^i = f(\tilde{P}_1^{''i}, \tilde{P}_{i+1}^{'n})$
- $a_4^n = a_3$

Lemma 2 implies that  $a_4^i \in \{a_1, a_2, a_3\}$  for every  $i$ . Therefore let  $j$  be the first integer such that  $a_4^j \neq a_1$ . One of two things results:

- $a_4^j = a_2$

but  $a_1P_j''a$ , since  $a_2$  is only in third position in  $P_j''$ . Therefore  $f(\tilde{P}_1^{''j-1}, \tilde{P}_j^{'n})P_j''f(\tilde{P}_1^{''j}, \tilde{P}_{j+1}^{'n})$ , so  $f$  is manipulable.

- $a_4^j = a_3$

Now, if  $a_1P_j'a_3$ , we also have  $a_1P_j''a_3$ . Therefore  $f(\tilde{P}_1^{''j-1}, \tilde{P}_j^{'n})P_j''f(\tilde{P}_1^{''j}, \tilde{P}_{j+1}^{'n})$ , and  $f$  is manipulable. If  $a_3P_j'a_1$ , we directly get  $f(\tilde{P}_1^{''j}, \tilde{P}_{j+1}^{'n})P_j'f(\tilde{P}_1^{''j-1}, \tilde{P}_j^{'n})$ , and  $f$  is still manipulable. We are led therefore to a contradiction in every case, which shows that  $F(P)$  is transitive.

Since  $F$  verifies all of Arrow's conditions,  $F$  must be dictatorial; let  $i$  be the dictator. Let  $P$  be any profile and arrange the choices in such a way that  $a_1P_ia_2P_i \dots$ . Since  $i$  is the

dictator, more precisely we have  $a_1 F(P) a_2$  and therefore  $f(\tilde{P}) = a_1$ . But, by construction,  $\tilde{P}$  coincides with  $P$  and  $f(P)$  is therefore  $a_1$ , the preferred choice of  $i$ , which concludes the proof showing that  $i$  is also a dictator for  $f$ .

The readers who are interested in the other ways of proofs are referred to Mas-Colell, Whinston, and Green (1995).

## Reference

- Arrow, K and G. Debreu, "Existence of Equilibrium for a Competitive Economy," *Econometrica*, 1954, 22.
- Arrow, K. J. and Hahn, F. H. (1971), *General Competitive Analysis*, (San Francisco, Holden Day).
- Debreu, G., *Theory of Value*, (Wiley, New York), 1959.
- Debreu, G. and H. Scarf, "A Limit Theorem on the Core of an Economy," *International Economic Review*, 4 (1963), 235-246.
- Gibbard, A., "Manipulation of Voting Schemes," *Econometrica*, 41 (1973), 587-601.
- Hildenbrand, W., A. P. Kirman, *Equilibrium Analysis: Variations on the Thems by Edgeworth and Walras*, North-Holland: New York, 1988.
- Jehle, G. A., and P. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, 1998, Chapters 7-8.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 10.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995, Chapters 18.
- Rawls, J., *A Theory of Justice*, Cambridge: Harvard University Press, 1971.
- Rubinstein Ariel, *Lecture Notes in Microeconomics (modeling the economic agent)*, Princeton University Press, 2005.



- Satterthwaite, M. A., "Strategy-Proofness and Arrow's Existence and Correspondences for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, 10 (1975), 187-217.
- Tian, G., "On the Constrained Walrasian and Lindahl Correspondences," *Economics Letters*, 26 (1988), pp. 299-303.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 21.
- Zhou, L., "Strictly Fair Allocations in Large Exchange Economies," *Journal of Economic Theory*, 57 (1992), 158-175

# Chapter 10

## General Equilibrium Under Uncertainty

### 10.1 Introduction

In this chapter, we apply the general equilibrium framework developed in Chapters 7 to 9 to economic situations involving the exchange and allocation of resources under conditions of uncertainty.

We begin by formalizing uncertainty by means of *states of the world* and then introduce the key idea of a *contingent commodity*: a commodity whose delivery is conditional on the realized state of the world. We then use these tools to define the concept of an *Arrow-Debreu equilibrium*. This is simply a Walrasian equilibrium in which contingent commodities are traded. It follows from the general theory of Chapter 8 that an Arrow-Debreu equilibrium results in a Pareto optimal allocation of risk.

In Section 10.4, we provide an important reinterpretation of the concept of Arrow-Debreu equilibrium. We show that, under the assumptions of *self-fulfilling*, *irrational expectations*, Arrow-Debreu equilibria can be implemented by combining trade in a certain restricted set of contingent commodities with *spot trade* that occurs *after* the resolution of uncertainty. This results in a significant reduction in the number of ex ante (i.e., before uncertainty) markets that must operate.

In Section 10.5 we briefly illustrate some of the welfare difficulties raised by the possibility of *incomplete markets*, that is, by the possibility of there being too few asset

markets to guarantee a fully Pareto optimal allocation of risk.

## 10.2 A Market Economy with Contingent Commodities

As in our previous chapters, we contemplate an environment with  $L$  physical commodities,  $n$  consumers, and  $J$  firms. The new element is that technologies, endowments, and preferences are now *uncertain*.

Throughout this chapter, we represent uncertainty by assuming that technologies, endowments, and preferences depend on the *state of the world*. A state of the world is to be understood as a complete description of a possible outcome of uncertainty. For simplicity we take  $S$  to be a finite set with (abusing notation slightly)  $S$  elements. A typical element is denoted  $s = 1, \dots, S$ .

We state the key concepts of a *(state-)contingent commodity* and a *(state-)contingent commodity vector*. Using these concepts we shall then be able to express the dependence of technologies, endowments, and preferences on the realized states of the world.

**Definition 10.2.1** *For every physical commodity  $\ell = 1, \dots, L$  and state  $s = 1, \dots, S$ , a unit of (state-)contingent commodity  $\ell s$  is a title to receive a unit of the physical good  $\ell$  if and only if  $s$  occurs. Accordingly, a (state-)contingent commodity vector is specified by*

$$x = (x_{11}, \dots, x_{L1}, \dots, x_{1S}, \dots, x_{LS}) \in \mathbb{R}^{LS},$$

*and is understood as an entitlement to receive the commodity vector  $(x_{1s}, \dots, x_{Ls})$  if state  $s$  occurs.*

We can also view a contingent commodity vector as a collection of  $L$  random variables, the  $l$ th random variable being  $(x_{l1}, \dots, x_{lS})$ .

With the help of the concept of contingent commodity vectors, we can now describe how the characteristics of economic agents depend on the state of the world. To begin, we let the endowments of consumer  $i = 1, \dots, n$  be a contingent commodity vector:

$$w_i = (w_{11i}, \dots, w_{L1i}, \dots, w_{1Si}, \dots, w_{LSi}) \in \mathbb{R}^{LS}.$$

The meaning of this is that if state  $s$  occurs then consumer  $i$  has endowment vector  $(w_{1si}, \dots, w_{Lsi}) \in \mathbb{R}^L$ .

The preferences of consumer  $i$  may also depend on the state of the world (e.g., the consumers enjoyment of wine may well depend on the state of his health). We represent this dependence formally by defining the consumers preferences over contingent commodity vectors. That is, we let the preferences of consumer  $i$  be specified by a rational preference relation  $\succsim_i$ , defined on a consumption set  $X_i \subset \mathbb{R}^{LS}$ .

The consumer evaluates contingent commodity vectors by first assigning to state  $s$  a probability  $\pi_{si}$  (which could have an objective or a subjective character), then evaluating the physical commodity vectors at state  $s$  according to a Bernoulli state-dependent utility function  $u_{si}(x_{1si}, \dots, x_{Lsi})$ , and finally computing the expected utility. That is, the preferences of consumer  $i$  over two contingent commodity vectors  $x_i, x'_i \in X_i \subset \mathbb{R}^{LS}$  satisfy

$$x_i \succsim_i x'_i \quad \text{if and only if} \quad \sum_s \pi_{si} u_{si}(x_{1si}, \dots, x_{Lsi}) \geq \sum_s \pi_{si} u_{si}(x'_{1si}, \dots, x'_{Lsi}).$$

It should be emphasized that the preferences  $\succsim_i$  are in the nature of ex ante preferences: the random variables describing possible consumptions are evaluated before the resolution of uncertainty.

Similarly, the technological possibilities of firm  $j$  are represented by a production set  $Y_j \subset \mathbb{R}^{LS}$ . The interpretation is that a *(state-)contingent production plan*  $y_j \in \mathbb{R}^{LS}$  is a member of  $Y_j$  if for every  $s$  the input-output vector  $(y_{1sj}, \dots, y_{Lsj})$  of physical commodities is feasible for firm  $j$  when state  $s$  occurs.

**Example 10.2.1** Suppose there are two states,  $s_1$  and  $s_2$ , representing good and bad weather. There are two physical commodities: seeds ( $\ell = 1$ ) and crops ( $\ell = 2$ ). In this case, the elements of  $Y_j$  are four-dimensional vectors. Assume that seeds must be planted before the resolution of the uncertainty about the weather and that a unit of seeds produces a unit of crops if and only if the weather is good. Then

$$y_j = (y_{11j}, y_{21j}, y_{12j}, y_{22j}) = (-1, 1, -1, 0)$$

is a feasible plan. Note that since the weather is unknown when the seeds are planted, the plan  $(-1, 1, 0, 0)$  is not feasible: the seeds, if planted, are planted in both states. Thus, in this manner we can imbed into the structure of  $Y_j$  constraints on production related to the timing of the resolution of uncertainty.

To complete the description of a private market economy it only remains to specify ownership shares for every consumer  $i$  and firm  $j$ . In principle, these shares could also be state-contingent. It will be simpler, however, to let  $\theta_{ij} \geq 0$  be the share of firm  $j$  owned by consumer  $i$  whatever the state. Of course  $\sum_j \theta_{ij} = 1$  for every  $i$ .

A private market economy with uncertain thus can be written

$$e = \left( S, \{X_i, w_i, \succsim_i\}_{i=1}^n, \{Y_j\}_{j=1}^J, \{\theta_{ij}\}_{i,j=1}^{n,J} \right).$$

## Information and the Resolution of Uncertainty

In the setting just described, time plays no explicit formal role. In reality, however, states of the world unfold over time. Figure 10.1 captures the simplest example. In the figure, we have a period 0 in which there is no information whatsoever on the true state of the world and a period 1 in which this information has been completely revealed.

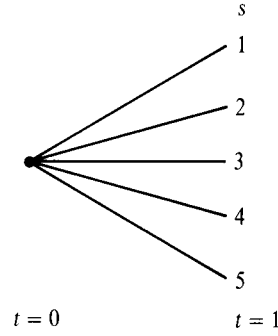


Figure 10.1: Two periods. Perfect information at  $t = 1$ .

The same methodology can be used to incorporate into the formalism a much more general temporal structure. Suppose we have  $T + 1$  dates  $t = 0, 1, \dots, T$  and, as before,  $S$  states, but assume that the states emerge gradually through a *tree*, as in Figure 10.2. Here final nodes stand for the possible states realized by time  $t = T$ , that is, for complete histories of the uncertain environment. When the path through the tree coincides for two states,  $s$  and  $s'$ , up to time  $t$ , this means that in all periods up to and including period  $t$ ,  $s$  and  $s'$  cannot be distinguished.

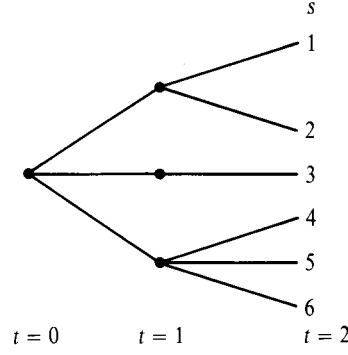


Figure 10.2: An information tree: gradual release of information.

Subsets of  $S$  are called *events*. A collection of events  $\mathcal{L}$  is an *information structure* if it is a partition, that is, if for every state  $s$  there is  $E \in \mathcal{L}$  with  $s \in E$  and for any two  $E, E' \in \mathcal{L}, E \neq E'$ , we have  $E \cap E' = \emptyset$ . The interpretation is that if  $s$  and  $s'$  belong to the same event in  $\mathcal{L}$  then  $s$  and  $s'$  cannot be distinguished in the information structure  $\mathcal{L}$ .

To capture formally a situation with sequential revelation of information we look at a *family* of information structures:  $(\mathcal{L}_0, \dots, \mathcal{L}_t, \dots, \mathcal{L}_T)$ . The process of information revelation makes the  $\mathcal{L}_t$  increasingly fine: once one has information sufficient to distinguish between two states, the information is not forgotten.

**Example 10.2.2** Consider the tree in Figure 10.2. We have

$$\begin{aligned}\mathcal{L}_0 &= (\{1, 2, 3, 4, 5, 6\}), \\ \mathcal{L}_1 &= (\{1, 2\}, \{3\}, \{4, 5, 6\}) \\ \mathcal{L}_2 &= (\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}).\end{aligned}$$

The partitions could in principle be different across individuals. However, we shall assume that the information structure is the same for all consumers.

A pair  $(t, E)$  where  $t$  is a date and  $E \in \mathcal{L}_t$  is called a *date-event*. Date-events are associated with the nodes of the tree. Each date-event except the first has a *unique predecessor*, and each date-event not at the end of the tree has one or more *successors*.

With this temporal modeling it is now necessary to be explicit about the time at which a physical commodity is available. Suppose there is a number  $H$  of basic physical

commodities (bread, leisure, etc.). We will use the double index  $ht$  to indicate the time at which a commodity  $h$  is produced, appears as endowment, or is available for consumption. Then  $x_{hts}$  stands for an amount of the physical commodity  $h$  available at time  $t$  along the path of state  $s$ .

Fortunately, this multi-period model can be formally reduced to the timeless structure introduced above. To see this, we define a new set of  $L = H(T + 1)$  physical commodities, each of them being one of these double-indexed (i.e.,  $ht$ ) commodities. We then say that a vector  $z \in \mathbb{R}^{LS}$  is *measurable* with respect to the family of information partitions  $(\mathcal{L}_0, \dots, \mathcal{L}_T)$  if, for every  $hts$  and  $hts'$ , we have that  $z_{hts} = z_{hts'}$  whenever  $s, s'$  belong to the same element of the partition  $\mathcal{L}_t$ . That is, whenever  $s$  and  $s'$  cannot be distinguished at time  $t$ , the amounts assigned to the two states cannot be different. Finally, we impose on endowments  $w_i \in \mathbb{R}^{LS}$ , consumption sets  $X_i \subset \mathbb{R}^{LS}$  and production sets  $Y_j \subset \mathbb{R}^{LS}$  the restriction that all their elements be measurable with respect to the family of information partitions. With this, we have reduced the multi-period structure to our original formulation.

## 10.3 Arrow-Debreu Equilibrium

We postulate the existence of a market for every contingent commodity  $ls$ . These markets open *before* the resolution of uncertainty. The price of the commodity is denoted by  $p_{ls}$ . Notice that for this market to be well defined it is indispensable that all economic agents be able to recognize the occurrence of  $s$ . That is, information should be *symmetric* across economic agents.

Formally, the market economy just described is nothing but a particular case of the economies we have studied in previous chapters. We can, therefore, apply to our market economy the concept of Walrasian equilibrium and, with it, all the theory developed so far. When dealing with contingent commodities it is customary to call the Walrasian equilibrium an *Arrow-Debreu equilibrium*.

**Definition 10.3.1** An allocation

$$(x_1^*, \dots, x_I^*, y_1^*, \dots, Y_J^*) \in X_1 \times \dots \times X_I \times Y_1 \times \dots \times Y_J \subset \mathbb{R}^{LS(I+J)}$$

and a system of prices for the contingent commodities  $p = (p_{11}, \dots, p_{LS}) \in \mathbb{R}^{LS}$  constitute an *Arrow-Debreu equilibrium* if:

- (i) For every  $j$ ,  $Y_j^*$  satisfies  $p \cdot y_j^* \geq p \cdot y_j$  for all  $y_j \in Y_j$ .
- (ii) For every  $i$ ,  $x_i^*$  is maximal for  $\succsim_i$  in the budget set  $\{x_i \in X_i: p \cdot x_i \leq p \cdot w_i + \sum_j \theta_{ij} p \cdot y_j^*\}$ .
- (iii)  $\sum_i x_i^* = \sum_j y_j^* + \sum_i w_i$ .

The positive and welfare theorems of Chapters 7 and 8 apply without modification to the Arrow-Debreu equilibrium. Recall that, in the present context, the convexity assumption takes on an interpretation in terms of risk aversion. For example, in the expected utility setting, the preference relation  $\succsim_i$  is convex if the Bernoulli utilities  $u_{si}(x_{si})$  are concave.

The Pareto optimality implication of Arrow-Debreu equilibrium says, effectively, that the possibility of trading in contingent commodities leads, at equilibrium, to an efficient allocation of risk.

It is important to realize that at any production plan the profit of a firm,  $p \cdot y_j$ , is a nonrandom amount of dollars. Productions and deliveries of goods do, of course, depend on the state of the world, but the firm is active in all the contingent markets and manages, so to speak, to insure completely. This has important implications for the justification of profit maximization as the objective of the firm.

**Example 10.3.1** Consider an exchange economy with  $n = 2, L = 1$ , and  $S = 2$ . This lends itself to an Edgeworth box representation because there are precisely two contingent commodities. In Figures xx(a) and xx(b) we have  $w_1 = (1, 0), w_2 = (0, 1)$ , and utility functions of the form  $\pi_{1i}u_i(x_{1i}) + \pi_{2i}u_i(x_{2i})$ , where  $(\pi_{1i}, \pi_{2i})$  are the subjective probabilities of consumer  $i$  for the two states. Since  $w_1 + w_2 = (1, 1)$  there is no aggregate uncertainty, and the state of the world determines only which consumer receives the endowment of the consumption good. Recall that for this model [in which the  $u_i(\cdot)$  do not depend on  $s$ ], the marginal rate of substitution of consumer  $i$  at any point where the consumption is the same in the two states equals the probability ratio  $\pi_{1i}/\pi_{2i}$ .

In Figure 10.3(a) the subjective probabilities are the same for the two consumers (i.e.,  $\pi_{11} = \pi_{12}$ ) and therefore the Pareto set coincides with the diagonal of the box (the box is a square and so the diagonal coincides with the 45-degree line, where the marginal rates of substitution for the two consumers are equal:  $\pi_{11}/\pi_{21} = \pi_{12}/\pi_{22}$ ). Hence, at equilibrium,



the two consumers insure completely; that is, consumer  $i$ 's equilibrium consumption does not vary across the two states. In Figure 10.3(b) the consumers subjective probabilities are different. In particular,  $\pi_{11} < \pi_{12}$  (i.e., the second consumer gives more probability to state 1). In this case, each consumer's equilibrium consumption is higher in the state he thinks comparatively more likely (relative to the beliefs of the other consumer).

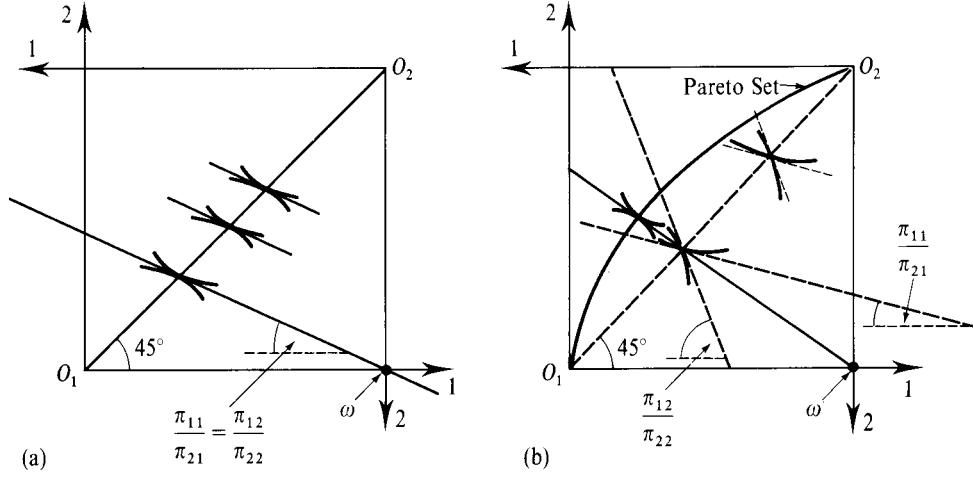


Figure 10.3: (a) No aggregate risk: some probability assessments. (b) No aggregate risk: different probability assessments.

## 10.4 Sequential Trade

The Arrow-Debreu framework provides a remarkable illustration of the power of general equilibrium theory. Yet, it is hardly realistic. Indeed, at an Arrow-Debreu equilibrium all trade takes place simultaneously and before the uncertainty is resolved. Trade is a one-shot affair. In reality, however, trade takes place to a large extent sequentially over time, and frequently as a consequence of information disclosures. The aim of this section is to introduce a first model of sequential trade and show that Arrow-Debreu equilibria can be reinterpreted by means of trading processes that actually unfold through time.

To be as simple as possible we consider only exchange economies. In addition, we take  $X_i = \mathbb{R}_+^{LS}$  for every  $i$ . To begin with, we assume that there are two dates,  $t = 0$  and  $t = 1$ , that there is no information whatsoever at  $t = 0$ , and that the uncertainty has

resolved completely at  $t = 1$ . Thus, the date-event tree is as in Figure xx.B.1. Again for simplicity, we assume that there is no consumption at  $t = 0$ .

Suppose that markets for the  $LS$  possible contingent commodities are set up at  $t = 0$ , and that  $(x_1^*, \dots, x_n^*) \in \mathbb{R}^{LSn}$  is an Arrow-Debreu equilibrium allocation with prices  $(p_{11}, \dots, p_{LS}) \in \mathbb{R}^{LS}$ . Recall that these markets are for delivery of goods at  $t = 1$  (they are commonly called *forward markets*). When period  $t = 1$  arrives, a state of the world  $s$  is revealed, contracts are executed, and every consumer  $i$  receives  $x_{si}^* = (x_{1si}^*, \dots, x_{Lsi}^*) \in \mathbb{R}^L$ . Imagine now that, after this but before the actual consumption of  $x_{si}^*$ , markets for the  $L$  physical goods were to open at  $t = 1$  (these are called *spot markets*). Would there be any incentive to trade in these markets? The answer is “no.” To see why, suppose that there were potential gains from trade among the consumers. that is, that there were  $x_{si} = (x_{1si}, \dots, x_{Lsi}) \in \mathbb{R}^L$  for  $i = 1, \dots, n$ , such that  $\sum_i x_{si} \leq \sum_i w_{si}$  and  $(x_{1i}^*, \dots, x_{si}, \dots, x_{si}^*) \succsim_i (x_{1i}^*, \dots, x_{si}^*, \dots, x_{si}^*)$  for all  $i$ , with at least one preference strict. It then follows from the definition of Pareto optimality that the Arrow-Debreu equilibrium allocation  $(x_1^*, \dots, x_n^*) \in \mathbb{R}^{LSn}$  is not Pareto optimal, contradicting the conclusion of the first welfare theorem. In summary at  $t = 0$  the consumers can trade directly to an overall Pareto optimal allocation; hence there is no reason for further trade to take place. In other words, ex ante pareto optimality implies ex post Pareto optimality and thus no ex post trade.

Matters are different if not all the  $LS$  contingent commodity markets are available at  $t = 0$ . Then the initial trade to a Pareto optimal allocation may not be feasible and it is quite possible that ex post (i.e., after the revelation of the state  $s$ ) the resulting consumption allocation is not Pareto optimal. There would then be an incentive to reopen the markets and retrade.

A most interesting possibility, first observed by Arrow (1953), is that, even if not all the contingent commodities are available at  $t = 0$ , it may still be the case under some conditions that the retrading possibilities at  $t = 1$  guarantee that Pareto optimality is reached, nevertheless. That is, the possibility of ex post trade can make up for an absence of some ex ante markets. In what follows, we shall verify that this is the case whenever at least one physical commodity can be traded contingently at  $t = 0$  if, in addition, spot markets occur at  $t = 1$  and the spot equilibrium prices are correctly anticipated at  $t = 0$ .

The intuition for this result is reasonably straightforward: if spot trade can occur within each state, then the only task remaining at  $t = 0$  is to transfer the consumer's overall purchasing power efficiently across states. This can be accomplished using contingent trade in a single commodity. By such a procedure we are able to reduce the number of required forward markets for  $LS$  to  $S$ .

Let us be more specific. At  $t = 0$  consumers have *expectations* regarding the spot prices prevailing at  $t = 1$  for each possible state  $s \in S$ . Denote the price vector expected to prevail in state  $s$  spot market by  $p_s \in \mathbb{R}^L$ , and the overall expectation vector by  $p = (p_1, \dots, p_S) \in \mathbb{R}^{LS}$ . Suppose that, in addition, at date  $t = 0$  there is trade in the  $S$  contingent commodities denoted by  $11$  to  $1S$ ; that is, there is contingent trade only in the physical good with the label 1. We denote the vector of prices for these contingent commodities traded at  $t = 0$  by  $q = (q_1, \dots, q_S) \in \mathbb{R}^S$ .

Faced with prices  $q \in \mathbb{R}^S$  at  $t = 0$  and expected spot prices  $(p_1, \dots, p_S) \in \mathbb{R}^{LS}$  at  $t = 1$ , every consumer  $i$  formulates a consumption, or trading, plan  $(z_{1i}, \dots, z_{Si}) \in \mathbb{R}^S$  for contingent commodities at  $t = 0$ , as well as a set of spot market consumption plans  $(x_{1i}, \dots, x_{Si}) \in \mathbb{R}^{LS}$  for the different states that may occur at  $t = 1$ . Of course, these plans must satisfy a budget constraint. Let  $U_i(\cdot)$  be a utility function for  $\succsim_i$ . Then the problem of consumer  $i$  can be expressed formally as

$$\begin{aligned}
 & \underset{\substack{(x_{1i}, \dots, x_{Si}) \in \mathbb{R}_+^{LS} \\ (z_{1i}, \dots, z_{Si}) \in \mathbb{R}^S}}}{Max} && U_i(x_{1i}, \dots, x_{Si}) \\
 & s.t. && (i) \sum_s q_s z_{si} \leq 0, \\
 & && (ii) p_s \cdot x_{si} \leq p_s \cdot w_{si} + p_{1s} z_{si} \text{ for every } s.
 \end{aligned} \tag{10.1}$$

Restriction (i) is the budget constraint corresponding to trade at  $t = 0$ . The family of restrictions (ii) are the budget constraints for the different spot markets. Note that the value of wealth at a state  $s$  is composed of two parts: the market value of the initial endowments,  $p_s \cdot w_{si}$ , and the market value of the amounts  $z_{si}$ , of good 1 bought or sold forward at  $t = 0$ . Observe that we are not imposing any restriction on the sign or the magnitude of  $z_{si}$ . If  $z_{si} < -w_{1si}$  then one says that at  $t = 0$  consumer  $i$  is selling good 1 *short*. This is because he is selling at  $t = 0$ , contingent on state  $s$  occurring, more than he has at  $t = 1$  if  $s$  occurs. Hence, if  $s$  occurs he will actually have to buy in the spot market the extra amount of the first good required for the fulfillment of his commitments. The

possibility of selling short is, however, indirectly limited by the fact that consumption, and therefore ex post wealth, must be nonnegative for every  $s$ .<sup>1</sup>

To define an appropriate notion of sequential trade we shall impose a key condition: Consumers' expectations must be *self-fulfilled*, or *rational*; that is, we require that consumers' expectations of the prices that will clear the spot markets for the different states  $s$  do actually clear them once date  $t = 1$  has arrived and a state  $s$  is revealed.

**Definition 10.4.1** A collection formed by a price vector  $q = (q_1, \dots, q_S) \in \mathbb{R}^S$  for contingent first good commodities at  $t = 0$ , a spot price vector

$$p_s = (p_{1s}, \dots, p_{LS}) \in \mathbb{R}^L$$

for every  $s$ , and, for every consumer  $i$ , consumption plans  $z_i^* = (z_{1i}^*, \dots, z_{Si}^*) \in \mathbb{R}^S$  at  $t = 0$  and  $x_i^* = (x_{1i}^*, \dots, x_{Si}^*) \in \mathbb{R}^{LS}$  at  $t = 1$  constitutes a *Radner equilibrium* [see Radner (1982)] if:

- (i) For every  $i$ , the consumption plans  $z_i^*, x_i^*$  solve problem (10.1).
- (ii)  $\sum_i z_{si}^* \leq 0$  and  $\sum_i x_{si}^* \leq \sum_i w_{si}$  for every  $s$ .

At a Radner equilibrium, trade takes place through time and, in contrast to the Arrow-Debreu setting, economic agents face a *sequence of budget sets*, one at each date-state (more generally, at every date-event).

We can see from an examination of problem (10.1) that all the budget constraints are homogeneous of degree zero with respect to prices. It is natural to choose the first commodity and to put  $p_{1s} = 1$  for every  $s$ , so that a unit of the  $s$  contingent commodity then pays off 1 dollar in state  $s$ . Note that this still leaves one degree of freedom, that corresponding to the forward trades at date 0 (so we could put  $q_1 = 1$ , or perhaps  $\sum_s q_s = 1$ ).

---

<sup>1</sup>Observe also that we have taken the wealth at  $t = 0$  to be zero (that is, there are no initial endowments of the contingent commodities). This is simply a convention. Suppose, for example, that we regard  $w_{1si}$ , the amount of good 1 available at  $t = 1$  in state  $s$ , as the amount of the  $s$  contingent commodity that  $i$  owns at  $t = 0$  (to avoid double counting, the initial endowment of commodity 1 in the spot market  $s$  at  $z = 1$  should simultaneously be put to zero). The budget constraints are then: (i)  $\sum_s q_s (z'_{si} - w_{1si}) \leq 0$  and (ii)  $p_s \cdot x_{si} \leq \sum_{l \neq 1} p_{ls} w_{l1} + p_{1s} z'_{si}$ , for every  $s$ . But letting  $z'_i = z_{si} + w_{1si}$ , we see that these are exactly the constraints of (19.D.1).

In the following proposition, which is the key result of this section, we show that for this model the set of Arrow-Debreu equilibrium allocations (induced by the arrangement of one-shot trade in  $LS$  contingent commodities) and the set of Radner equilibrium allocations (induced by contingent trade in only one commodity, sequentially followed by spot trade) are identical.

**Proposition 10.4.1** *We have:*

- (i) *If the allocation  $x^* \in \mathbb{R}^{LSn}$  and the contingent commodities price vector  $(p_1, \dots, p_S) \in \mathbb{R}_{++}^{LS}$  constitute an Arrow-Debreu equilibrium, then there are prices  $q \in \mathbb{R}_{++}^S$  for contingent first good commodities and consumption plans for these commodities  $z^* = (z_1^*, \dots, z_n^*) \in \mathbb{R}^{Sn}$  such that the consumptions plans  $x^*, z^*$ , the prices  $q$ , and the spot prices  $(p_1, \dots, p_S)$  constitute a Radner equilibrium.*
- (ii) *Conversely, if the consumption plans  $x^* \in \mathbb{R}^{LSn}, z^* \in \mathbb{R}^{Sn}$  and prices  $q \in \mathbb{R}_{++}^S, (p_1, \dots, p_S) \in \mathbb{R}_{++}^{LS}$  constitute a Radner equilibrium, then there are multipliers  $(\mu_1, \dots, \mu_S) \in \mathbb{R}_{++}^S$  such that the allocation  $x^*$  and the contingent commodities price vector  $(\mu_1 p_1, \dots, \mu_S p_S) \in \mathbb{R}_{++}^{LS}$  constitute an Arrow-Debreu equilibrium. (The multiplier  $\mu_s$  is interpreted as the value, at  $t = 0$ , of a dollar at  $t = 1$  and state  $s$ .)*

PROOF:

- i. It is natural to let  $q_s = p_{1s}$  for every  $s$ . With this we claim that, for every consumer  $i$ , the budget set of the Arrow-Debreu problem,

$$B_i^{AD} = \{(x_{1i}, \dots, x_{Si}) \in \mathbb{R}_+^{LS} : \sum_s p_s \cdot (x_{si} - w_{si}) \leq 0\},$$

is identical to the budget set of the Radner problem,

$$\begin{aligned} B_i^{AD} &= \{(x_{1i}, \dots, x_{Si}) \in \mathbb{R}_+^{LS} : \text{there are } (z_{1i}, \dots, z_{Si}) \\ &\text{such that } \sum_s p_s z_{si} \leq 0 \text{ and } p_s \cdot (x_{si} - w_{si}) \leq p_{1s} z_{si} \text{ for every } s\}. \end{aligned}$$

To see this, suppose that  $x_i = (x_{1i}, \dots, x_{Si}) \in B_i^{AD}$ . For every  $s$ , denote  $z_{si} = (1/p_{1s})p_s \cdot (x_{si} - w_{si})$ . Then  $\sum_s p_s z_{si} = \sum_s p_{1s} z_{si} = \sum_s p_s \cdot (x_{si} - w_{si}) \leq 0$  and

$p_s \cdot (x_{si} - w_{si}) = p_1 z_{si}$  for every  $s$ . Hence,  $x_1 \in B_i^R$ . Conversely, suppose that  $x_i = (x_{1i}, \dots, x_{Si}) \in B_i^R$ ; that is, for some  $(z_{1i}, \dots, z_{Si})$  we have  $\sum_s q_s z_{si} \leq 0$  and  $p_s \cdot (x_{si} - w_{si}) = p_1 z_{si}$ , for every  $s$ . Summing over  $s$  gives  $\sum_s p_s \cdot (x_{si} - w_{si}) \leq \sum_s p_1 z_{si} = \sum_s q_s z_{si} \leq 0$ . Hence,  $x_i \in B^{AD}_i$ .

We conclude that our Arrow-Debreu equilibrium allocation is also a Radner equilibrium allocation supported by  $q = (p_{11}, \dots, p_{1S}) \in \mathbb{R}^S$ , the spot prices  $(p_1, \dots, p_S)$ , and the contingent trades  $(z_{1i}^*, \dots, z_{Si}^*) \in \mathbb{R}^S$  defined by  $z_{si}^* = (1/p_{1s})p_s \cdot (x_{si}^* - w_{si})$ . Note that the contingent markets clear since, for every  $s$ ,  $\sum_i z_{si}^* = (1/p_{1s})p_s \cdot [\sum_i (x_{si}^* - w_{si})] \leq 0$ .

- ii. Choose  $\mu_s$  so that  $\mu_s p_{1s} = q_s$ . Then we can rewrite the Radner budget set of every consumer  $i$  as  $B_i^R = \{(x_{1i}, \dots, x_{Si}) \in \mathbb{R}^{LS} : \text{there are } (z_{1i}, \dots, z_{Si}) \text{ such that } \sum_s q_s z_{si} \leq 0 \text{ and } \mu_s p_s \cdot (x_{si} - w_{si}) \leq q_s z_{si} \text{ for every } s\}$ . But from this we can proceed as we did in part (i) and rewrite the constraints, and therefore the budget set, in the Arrow-Debreu form:

$$B_i^r = B_i^{AD} = \{(x_{1i}, \dots, x_{Si}) \in \mathbb{R}^{LS} : \sum_s \mu_s p_s \cdot (x_{si} - w_{si}) \leq 0\}.$$

Hence, the consumption plan  $x_i^*$  is also preference maximizing in the budget set  $B_i^{AD}$ . Since this is true for every consumer  $i$ , we conclude that the price vector  $(\mu_s p_1, \dots, \mu_S p_S) \in \mathbb{R}^{LS}$  clears the markets for the  $LS$  contingent commodities.

**Example 10.4.1 :** Consider a two-good, two-state, two-consumer pure exchange economy. Suppose that the two states are equally likely and that every consumer has the same, state-independent, Bernoulli utility function  $u(x_{si})$ . The consumers differ only in their initial endowments. The aggregate endowment vectors in the two states are the same; however, endowments are distributed so that consumer 1 gets everything in state 1 and consumer 2 gets everything in state 2. (See Figure 10.4.)

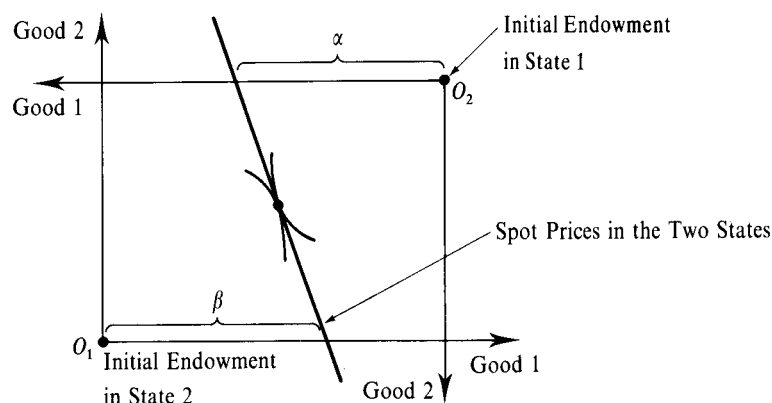


Figure 10.4: Reaching the Arrow-Debreu equilibrium by means of contingent trade in the first good only.

By the symmetry of the problem, at an Arrow-Debreu equilibrium each consumer gets, in each state, half of the total endowment of each good. In Figure 19.D.1, we indicate how these consumptions will be reached by means of contingent trade in the first commodity and spot markets. The spot prices will be the same in the two states. The first consumer will sell an amount  $\alpha$  of the first good contingent on the occurrence of the first state and will in exchange buy an amount  $\beta$  of the same good contingent on the second state.

**Remark 10.4.1** It is important to emphasize that, although the concept of Radner equilibrium cuts down the number of contingent commodities required to attain optimality (from  $LS$  to  $S$ ), this reduction is not obtained free of charge. With the smaller number of forward contracts, the correct anticipation of future spot prices becomes crucial.

## 10.5 Incomplete Markets

In this section we explore the implications of having fewer than  $S$  assets, that is, of having an asset structure that is necessarily incomplete. We pursue this point in the two-period framework of the previous sections.

We begin by observing that when  $K < S$  a Radner equilibrium need not be Pareto optimal. This is not surprising: if the possibilities of transferring wealth across states are limited, then there may well be a welfare loss due to the inability to diversify risks to

the extent that would be convenient. Just consider the extreme case where there are no assets at all. The following example provides another interesting illustration of this type of failure.

**Example 10.5.1** [Sunspots]. Suppose that preferences admit an expected utility representation and that the set of states  $S$  is such that, first, the probability estimates for the different states are the same across consumers (i.e.,  $\pi_{si} = \pi'_{si} = \pi_s$  for all  $i, i'$ , and  $s$ ) and second, that the states do not affect the fundamentals of the economy; that is, the Bernoulli utility functions and the endowments of every consumer  $i$  are uniform across states [i.e.,  $u_{si}(\cdot) = u_i(\cdot)$  and  $w_{si} = w_i$ , for all  $s$ ]. Such a set of states is called a *sunspot* set. The question we shall address is whether in these circumstances the Radner equilibrium allocations can assign varying consumptions across states. An equilibrium where this happens is called a *sunspot equilibrium*.

Under the assumption that consumers are strictly risk averse, so that the utility functions  $u_i(\cdot)$  are strictly concave, *any Pareto optimal allocation*  $(x_1, \dots, x_n) \in \mathbb{R}_+^{LSn}$  *must be uniform across states* (or *state independent*); that is, for every  $i$  we must have  $x_{1i} = x_{2i} = \dots = x_{si} = \dots = x_{Si}$ . To see this, suppose that, for every  $i$  and  $s$ , we replace the consumption bundle of consumer  $i$  in state  $s$ ,  $x_{si} \in \mathbb{R}_+^L$ , by the expected consumption bundle of this consumer:  $\bar{x}_i = \sum_s \pi_s x_{si} \in \mathbb{R}_+^L$ . The new allocation is state independent, and it is also feasible because

$$\sum_i \bar{x}_i = \sum_i \sum_s \pi_s x_{si} = \sum_s \pi_s \left( \sum_i x_{si} \right) \leq \sum_s \pi_s \left( \sum_i w_i \right) = \sum_i w_i.$$

By the concavity of  $u_i(\cdot)$  it follows that no consumer is worse off:

$$\sum_s \pi_s u_i(\bar{x}_i) = u_i(\bar{x}_i) = u_i \left( \sum_s \pi_s x_{si} \right) \geq \sum_s \pi_s u_i(x_{si}) \text{ for every } i.$$

Because of the Pareto optimality of  $(x_1, \dots, x_n)$ , the above weak inequalities must in fact be equalities; that is,  $u_i(\bar{x}_i) = \sum_s \pi_s u_i(x_{si})$  for every  $i$ . But, if so, then the strict concavity of  $u_i(\cdot)$  yields  $x_{si} = \bar{x}_i$  for every  $s$ . In summary: the Pareto optimal allocation  $(x_1, \dots, x_n) \in \mathbb{R}^{LSn}$  is state independent.

From the state independence of Pareto optimal allocations and the first welfare theorem we reach the important conclusion that if a system of complete markets over the states  $S$  can be organized, then the equilibria are *sunspot free*, that is, consumption is



uniform across states. In effect, traders wish to insure completely and have instruments to do so.

It turns out, however, that if there is not a complete set of insurance opportunities, then the above conclusion does not hold true. Sunspot-free, Pareto optimal equilibria always exist (just make the market “not pay attention” to the sunspot. But it is now possible for the consumption allocation of some Radner equilibria to depend on the state, and consequently to fail the Pareto optimality test. In such an equilibrium consumers expect different prices in different states, and their expectations end up being self-fulfilling. The simplest, and most trivial, example is when there are no assets whatsoever ( $K = 0$ ). Then a system of spot prices  $(p_1, \dots, p_s) \in \mathbb{R}^{LS}$  is a Radner equilibrium if and only if every  $p_s$  is a Walrasian equilibrium price vector for the spot economy defined by  $\{(u_i(\cdot), w_i)\}_{i=1}^{i=n}$ . If, as is perfectly possible, this economy admits several distinct Walrasian equilibria, then by selecting different equilibrium price vectors for different states, we obtain a sunspot equilibrium, and hence a Pareto inefficient Radner equilibrium.

We have seen that Radner equilibrium allocations need not be Pareto optimal, and so, in principle, there may exist reallocations of consumption that make all consumers at least as well off, and at least one consumer strictly better off. It is important to recognize, however, that this *need not* imply that a welfare authority who is “as constrained in interstate transfers as the market is” can achieve a Pareto optimum. An allocation that cannot be Pareto improved by such an authority is called a *constrained Pareto optimum*. A more significant and reasonable welfare question to ask is, therefore, whether Radner equilibrium allocations are constrained Pareto optimal. We now address this matter. This is a typical instance of a *second-best* welfare issue.

To proceed with the analysis we need a precise description of the constrained feasible set and of the corresponding notion of constrained Pareto optimality. This is most simply done in the context where there is a single commodity per state, that is,  $L = 1$ . The important implication of this assumption is that then the amount of consumption good that any consumer  $i$  gets in the different states is entirely determined by the portfolio  $z_i$ . Indeed,  $x_{si} = \sum_k z_{ki} r_{sk} + w_{si}$ . Hence, we can let

$$U_i^*(z_i) = U_i^*(z_{1i}, \dots, z_{Ki}) = U_i\left(\sum_k z_{ki} r_{1k} + w_{1i}, \dots, \sum_k z_{ki} r_{sk} + w_{si}\right)$$

denote the utility induced by the portfolio  $z_i$ . The definition of constrained Pareto optimality is then quite natural.

**Definition 10.5.1** The asset allocation  $(z_1, \dots, z_n) \in \mathbb{R}^{Kn}$  is constrained Pareto optimal if it is feasible (i.e.,  $\sum_i z_i \leq 0$ ) and if there is no other feasible asset allocation  $(z'_1, \dots, z'_n) \in \mathbb{R}^{Kn}$  such that

$$U_i^*(z'_1, \dots, z'_n) \geq U_i^*(z_1, \dots, z_n) \text{ for every } i,$$

with at least one inequality strict.

In this  $L = 1$  context the utility maximization problem of consumer  $i$  becomes

$$\begin{aligned} \underset{z_i \in \mathbb{R}^K}{Max} \quad & U_i^*(z_{1i}, \dots, z_{Ki}) \\ \text{s.t.} \quad & q \cdot z_i \leq 0 . \end{aligned}$$

Suppose that  $z_i \in \mathbb{R}^K$  for  $i = 1, \dots, n$ , is a family of solutions to these individual problems, for the asset price vector  $q \in \mathbb{R}^K$ . Then  $q \in \mathbb{R}^K$  is a Radner equilibrium price if and only if  $\sum_i z_i^* \leq 0$ .<sup>2</sup> To it we can apply the first welfare theorem and reach the conclusion of the following proposition.

**Proposition 10.5.1** *Suppose that there two periods and only one consumption good in the second period. Then any Radner equilibrium is constrained Pareto optimal in the sense that there is no possible redistribution of assets in the first period that leaves every consumer as well off and at least one consumer strictly better off.*

The situation considered in Proposition 10.5.1 is very particular in that once the initial asset portfolio of a consumer is determined, his overall consumption is fully determined: with only one consumption good, there are no possibilities for trade once the state occurs. In particular, second-period relative prices do not matter, simply because there are no such prices. Things change if there is more than one consumption good in the second period, or if there are more than two periods. such an example can be found in Mas-Colell, Whinston, and Green (1995, p. 711).

---

<sup>2</sup>Recall that, given  $z_i$  the consumptions in every state are determined. Also, the price of consumption good in every state is formally fixed to be 1.

## Reference

- Jehle, G. A., and P. Reny, *Advanced Microeconomic Theory*, Addison-Wesley, 1998, Chapters 7-8.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 10.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic Theory*, Oxford University Press, 1995, Chapters 18.
- Rubinstein Ariel, *Lecture Notes in Microeconomics (modeling the economic agent)*, Princeton University Press, 2005.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 21.

## Part IV

# Externalities and Public Goods

In Chapters 7 and 8, we have introduced the notions of competitive equilibrium and Pareto optimality, respectively. The concept of competitive equilibrium provides us with an appropriate notion of market equilibrium for competitive market economies. The concept of Pareto optimality offers a minimal and uncontroversial test that any social optimal economic outcome should pass since it is a formulation of the idea that there is no further improvement in society, and it conveniently separates the issue of economic efficiency from more controversial (and political) questions regarding the ideal distribution of well-being across individuals.

The important results and insights obtained in Chapters 7 and 8 are the First and Second Fundamental Theorems of Welfare Economics. The first welfare theorem provides a set of conditions under which a market economy will achieve a Pareto optimal; it is, in a sense, the formal expression of Adam Smith's claim about the "invisible hand" of the market. The second welfare theorem goes even further. It states that under the same set of conditions as the first welfare theorem plus convexity and continuity conditions, all Pareto optimal outcomes can in principle be implemented through the market mechanism by appropriately redistributing wealth and then "letting the market work."

Thus, in an important sense, the general equilibrium theory establishes the perfectly competitive case as a benchmark for thinking about outcomes in market economies. In particular, any inefficiency that arise in a market economy, and hence any role for Pareto-improving market intervention, must be traceable to a violation of at least one of these assumptions of the first welfare theorem. The remainder of the notes, can be viewed as a development of this theme, and will study a number of ways in which actual markets may depart from this perfectly competitive ideal and where, as a result, market equilibria fail to be Pareto optimal, a situation known market failure.

In the current part, we will study externalities and public goods in Chapter 11 and Chapter 12, respectively. In both cases, the actions of one agent directly affect the utility or production of other agents in the economy. We will see these nonmarketed "goods" or "bads" lead to a non-Pareto optimal outcome in general; thus a market failure. It turns out that private markets are often not a very good mechanism in the presence of externalities and public goods. We will consider situations of incomplete information which also result in non-Pareto optimal outcomes in general in Part V.

It is greatly acknowledged that the materials are largely drawn from the references in the end of each charter, especially from Varian (1992).

# Chapter 11

## Externalities

### 11.1 Introduction

In this chapter we deal with the case of externalities so that a market equilibrium may lead to non-Pareto efficient allocations in general, and thus there is a market failure. The reason is that there are things that people care about are not priced. Externality can happen in both cases of consumption and production.

*Consumption Externality:*

$u_i(x_i)$  : without preference externality

$u_i(x_1, \dots, x_n)$  : with preference externality

in which other individuals' consumptions effect an individual's utility.

**Example 11.1.1** (i) One person's quiet environment is disturbed by another person's local stereo.

(ii) Mr. A hates Mr. D smoking next to him.

(ii) Mr. A's satisfaction decreases as Mr. C's consumption level increases, because Mr. A envies Mr. C's lifestyle.

*Production Externality:*

A firm's production includes arguments other than its own inputs.

For example, downstream fishing is adversely affected by pollutants emitted from an upstream chemical plant.

This leads to an examination of various suggestions for alternative ways to allocate resources that may lead to efficient outcomes. Achieving an efficient allocation in the presence of externalities essentially involves making sure that agents face the correct prices for their actions. Ways of solving externality problem include taxation, regulation, property rights, merges, etc.

## 11.2 Consumption Externalities

When there are no consumption externalities, agent  $i$ 's utility function is a function of only his own consumption:

$$u_i(x_i) \tag{11.1}$$

In this case, the first order conditions for the competitive equilibrium are given by

$$MRS_{xy}^A = \frac{p_x}{p_y} = MRS_{xy}^B$$

and the first order conditions for Pareto efficiency are given by:

$$MRS_{xy}^A = MRS_{xy}^B.$$

Thus, because of price-taking behavior, every competitive equilibrium implies Pareto efficiency if utility functions are quasi-concave.

The main purpose of this section is to show that a competitive equilibrium allocation is not in general Pareto efficient when there exists an externality in consumption. We show this by examining that the first order conditions for a competitive equilibrium is not in general the same as the first order conditions for Pareto efficient allocations in the presence of consumption externalities. The following materials are mainly absorbed from Tian and Yang (2009).

Consider the following simple two-person and two-good exchange economy.

$$u_A(x_A, x_B, y_A) \tag{11.2}$$

$$u_B(x_A, x_B, y_B) \tag{11.3}$$



which are assumed to be strictly increasing in his own goods consumption, quasi-concave, and satisfies the Inada condition  $\frac{\partial u}{\partial x_i}(0) = +\infty$  and  $\lim_{x_i \rightarrow 0} \frac{\partial u}{\partial x_i} x_i = 0$  so it results in interior solutions. We further assume that the gradient of  $u_i(\cdot)$  is nonzero at Pareto efficient allocations. Here good  $x$  results in consumption externalities.

The first order conditions for the competitive equilibrium are the same as before:

$$MRS_{xy}^A = \frac{p_x}{p_y} = MRS_{xy}^B.$$

We now find the first order conditions for Pareto efficient allocations in exchange economies with externalities. Thus Pareto efficient allocations  $x^*$  can be completely determined by the FOCs of the following problem.

$$\begin{aligned} & \underset{x \in \mathbb{R}_{++}^4}{Max} u_B(x_A, x_B, y_B) \\ & s.t. \ x_A + x_B \leq w_x \\ & \quad y_A + y_B \leq w_y \\ & \quad u_A(x_A, x_B, y_A) \geq u_A(x_A^*, x_B^*, y_A^*) \end{aligned}$$

The first order conditions are

$$x_A : \frac{\partial u_B}{\partial x_A} - \lambda_x + \mu \frac{\partial u_A}{\partial x_A} = 0 \quad (11.4)$$

$$y_A : -\lambda_y + \mu \frac{\partial u_A}{\partial y_A} = 0 \quad (11.5)$$

$$x_B : \frac{\partial u_B}{\partial x_B} - \lambda_x + \mu \frac{\partial u_A}{\partial x_B} = 0 \quad (11.6)$$

$$y_B : \frac{\partial u_B}{\partial y_B} - \lambda_y = 0 \quad (11.7)$$

$$\lambda_x : w_x - x_A - x_B \geq 0, \lambda_x \geq 0, \lambda_x (w_x - x_A - x_B) = 0 \quad (11.8)$$

$$\lambda_y : w_y - y_A - y_B \geq 0, \lambda_y \geq 0, \lambda_y (w_y - y_A - y_B) = 0 \quad (11.9)$$

$$\mu : u_A - u_A^* \geq 0, \mu \geq 0, \mu (u_A - u_A^*) = 0 \quad (11.10)$$

By (11.7),  $\lambda_y = \frac{\partial u_B}{\partial y_B} > 0$ , and thus by (11.9),

$$y_A + y_B = w_y \quad (11.11)$$

which means there is never destruction of the good which does not exhibit a negative externality. Also, by (11.5) and (11.7), we have

$$\mu = \frac{\frac{\partial u_B}{\partial y_B}}{\frac{\partial u_A}{\partial y_A}}. \quad (11.12)$$

Then, by (11.4) and (11.5), we have

$$\frac{\lambda_x}{\lambda_y} = \left[ \frac{\frac{\partial u_A}{\partial x_A}}{\frac{\partial u_A}{\partial y_A}} + \frac{\frac{\partial u_B}{\partial x_A}}{\frac{\partial u_B}{\partial y_B}} \right] \quad (11.13)$$

and by (11.6) and (11.7), we have

$$\frac{\lambda_x}{\lambda_y} = \left[ \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} + \frac{\frac{\partial u_A}{\partial x_B}}{\frac{\partial u_A}{\partial y_A}} \right] \quad (11.14)$$

Thus, by (11.13) and (11.14), we have

$$\frac{\frac{\partial u_A}{\partial x_A}}{\frac{\partial u_A}{\partial y_A}} + \frac{\frac{\partial u_B}{\partial x_A}}{\frac{\partial u_B}{\partial y_B}} = \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} + \frac{\frac{\partial u_A}{\partial x_B}}{\frac{\partial u_A}{\partial y_A}}, \quad (11.15)$$

which expresses the equality of the social marginal rates of substitution for the two consumers at Pareto efficient points. Thus, we immediately have the following conclusion:

A competitive equilibrium allocations may not be Pareto optimal because the first order conditions' for competitive equilibrium and Pareto optimality are not the same.

From the above marginal equality condition, we know that, in order to evaluate the relevant marginal rates of substitution for the optimality conditions, we must take into account both the direct and indirect effects of consumption activities in the presence of externalities. That is, to achieve Pareto optimality, when one consumer increases the consumption of good  $x$ , not only does the consumer's consumption of good  $y$  need to change, the other consumer's consumption of good  $y$  must also be changed. Therefore the social marginal rate of substitution of good  $x$  for good  $y$  by consumer  $i$  equals  $\frac{\frac{\partial u_i}{\partial x_i}}{\frac{\partial u_i}{\partial y_i}} + \frac{\frac{\partial u_j}{\partial x_i}}{\frac{\partial u_j}{\partial y_j}}$ .

Solving (11.4) and (11.6) for  $\mu$  and  $\lambda_x$ , we have

$$\mu = \frac{\frac{\partial u_B}{\partial x_B} - \frac{\partial u_B}{\partial x_A}}{\frac{\partial u_A}{\partial x_A} - \frac{\partial u_A}{\partial x_B}} > 0 \quad (11.16)$$

and

$$\lambda_x = \frac{\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} - \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A}}{\frac{\partial u_A}{\partial x_A} - \frac{\partial u_A}{\partial x_B}}. \quad (11.17)$$

When the consumption externality is positive, from (11.13) or (11.14), we can easily see that  $\lambda_x$  is always positive since  $\lambda_y = \frac{\partial u_B}{\partial y_B} > 0$ . Also, when no externality or a one-sided externality<sup>1</sup> exists, by either (11.13) or (11.14),  $\lambda_x$  is positive. Thus, the marginal equality condition (11.15) and the balanced conditions, completely determine all Pareto efficient allocations for these cases. However, when there are negative externalities for both consumers, the Kuhn-Tucker multiplier  $\lambda_x$  directly given by (11.17) or indirectly given by (11.13) or (11.14) is the sum of a negative and positive term, and thus the sign of  $\lambda_x$  may be indeterminate. However, unlike the claim in some textbooks such as Varian (1992, 438), the marginal equality condition, (11.15), and the balanced conditions may not guarantee finding Pareto efficient allocations correctly.

To guarantee that an allocation is Pareto efficient in the presence of negative externalities, we must require  $\lambda_x \geq 0$  at efficient points, which in turn requires that social marginal rates of substitution be nonnegative, that is,

$$\frac{\frac{\partial u_A}{\partial x_A}}{\frac{\partial u_A}{\partial y_A}} + \frac{\frac{\partial u_B}{\partial x_A}}{\frac{\partial u_B}{\partial y_B}} = \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} + \frac{\frac{\partial u_A}{\partial x_B}}{\frac{\partial u_A}{\partial y_A}} \geq 0, \quad (11.18)$$

or equivalently requires both (11.15) and

$$\underbrace{\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B}}_{\text{(joint marginal benefit)}} \geq \underbrace{\frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A}}_{\text{(joint marginal cost)}} \quad (11.19)$$

for all Pareto efficient points.

We can interpret the term in the left-hand side of (11.19),  $\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B}$ , as the joint marginal benefit of consuming good  $x$ , and the term in the right-hand side,  $\frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A}$ , as the joint marginal cost of consuming good  $x$  because the negative externality hurts the consumers. To consume the goods efficiently, a necessary condition is that the joint marginal benefit of consuming good  $x$  should not be less than the joint cost of consuming good  $x$ .

Thus, the following conditions

$$(PO) \left\{ \begin{array}{l} \frac{\frac{\partial u_A}{\partial x_A}}{\frac{\partial u_A}{\partial y_A}} + \frac{\frac{\partial u_B}{\partial x_A}}{\frac{\partial u_B}{\partial y_B}} = \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} + \frac{\frac{\partial u_A}{\partial x_B}}{\frac{\partial u_A}{\partial y_A}} \geq 0 \\ y_A + y_B = w_y \\ x_A + x_B \leq w_x \\ \left( \frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} - \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A} \right) (w_x - x_A - x_B) = 0 \end{array} \right.$$

---

<sup>1</sup>Only one consumer imposes an externality on another consumer.

constitute a system (PO) from which all Pareto efficient allocations can be obtained. We can do so by considering three cases.

Case 1. When  $\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} > \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A}$ , or equivalently  $\frac{\frac{\partial u_A}{\partial x_A}}{\frac{\partial u_A}{\partial y_A}} + \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} = \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} + \frac{\frac{\partial u_A}{\partial x_B}}{\frac{\partial u_A}{\partial y_A}} > 0$ ,  $\lambda_x > 0$  and thus the last two conditions in the above system (PO) reduce to  $x_A + x_B = w_x$ . In this case, there is no destruction. Substituting  $x_A + x_B = w_x$  and  $y_A + y_B = w_y$  into the marginal equality condition (11.15), it would give us a relationship between  $x_A$  and  $y_A$ , which exactly defines the Pareto efficient allocations.

Case 2. When the joint marginal benefit equals the joint marginal cost:

$$\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} = \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A}, \quad (11.20)$$

then

$$\frac{\frac{\partial u_A}{\partial x_A}}{\frac{\partial u_A}{\partial y_A}} + \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} = \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} + \frac{\frac{\partial u_A}{\partial x_B}}{\frac{\partial u_A}{\partial y_A}} = 0 \quad (11.21)$$

and thus  $\lambda_x = 0$ . In this case, when  $x_A + x_B \leq w_x$ , the necessity of destruction is indeterminant. However, even when destruction is necessary, we can still determine the set of Pareto efficient allocations by using  $y_A + y_B = w_y$  and the zero social marginal equality conditions (11.21). Indeed, after substituting  $y_A + y_B = w_y$  into (11.21), we can solve for  $x_A$  in terms  $y_A$ .

Case 3. When  $\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} < \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A}$  for any allocations that satisfy  $x_A + x_B = w_x$ ,  $y_A + y_B = w_y$ , and the marginal equality condition (11.15), the social marginal rates of substitution must be negative. Hence, the allocation will not be Pareto efficient. In this case, there must be a destruction for good  $x$  for Pareto efficiency, and a Pareto efficient allocation satisfies (11.21).

Summarizing, we have the following proposition that provides two categories of sufficiency conditions for characterizing whether or not there should be destruction of endowment  $w_x$  in achieving Pareto efficient allocations.

**Proposition 11.2.1** *For  $2 \times 2$  pure exchange economies, suppose that utility functions  $u_i(x_A, x_B, y_i)$  are continuously differentiable, strictly quasi-concave, and  $\frac{\partial u_i(x_A, x_B, y_i)}{\partial x_i} > 0$  for  $i = A, B$ .*

(1) *If the social marginal rates of substitution are positive at a Pareto effi-*

cient allocation  $x^*$ ,<sup>2</sup> then there is no destruction of  $w_x$  in achieving Pareto efficient allocation  $x^*$ .

(2) If the social marginal rates of substitution are negative for any allocation  $(x_A, x_B)$  satisfying  $x_A + x_B = w_x$ ,  $y_A + y_B = w_y$ , and the marginal equality condition (11.15), then there is destruction of  $w_x$  in achieving any Pareto efficient allocation  $x^*$ . That is,  $x_A^* + x_B^* < w_x$  and  $x^*$  is determined by  $y_A + y_B = w_y$  and (11.21).

Thus, from the above proposition, we know that a sufficient condition for destruction is that for any allocation  $(x_A, y_A, x_B, y_B)$ ,

$$\left. \begin{aligned} \frac{\frac{\partial u_A}{\partial x_A}}{\frac{\partial u_A}{\partial y_A}} + \frac{\frac{\partial u_B}{\partial x_A}}{\frac{\partial u_B}{\partial y_B}} &= \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} + \frac{\frac{\partial u_A}{\partial x_B}}{\frac{\partial u_A}{\partial y_A}} \\ x_A + x_B &= w_x \\ y_A + y_B &= w_y \end{aligned} \right\} \Rightarrow \frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} < \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A}. \quad (11.22)$$

A sufficient condition for all Pareto efficient allocations  $(x_A, y_A, x_B, y_B)$  for no destruction is,

$$\left. \begin{aligned} \frac{\frac{\partial u_A}{\partial x_A}}{\frac{\partial u_A}{\partial y_A}} + \frac{\frac{\partial u_B}{\partial x_A}}{\frac{\partial u_B}{\partial y_B}} &= \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} + \frac{\frac{\partial u_A}{\partial x_B}}{\frac{\partial u_A}{\partial y_A}} \\ x_A + x_B &\leq w_x \\ y_A + y_B &= w_y \end{aligned} \right\} \Rightarrow \frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} > \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A} \quad (11.23)$$

**Example 11.2.1** Consider the following utility function:

$$u_i(x_A, x_B, y_i) = \sqrt{x_i y_i} - x_j, \quad i \in \{A, B\}, j \in \{A, B\}, j \neq i$$

By the marginal equality condition (11.15), we have

$$\left( \sqrt{\frac{y_A}{x_A}} + 1 \right)^2 = \left( \sqrt{\frac{y_B}{x_B}} + 1 \right)^2 \quad (11.24)$$

and thus

$$\frac{y_A}{x_A} = \frac{y_B}{x_B}. \quad (11.25)$$

Let  $x_A + x_B \equiv \bar{x}$ . Substituting  $x_A + x_B = \bar{x}$  and  $y_A + y_B = w_y$  into (11.25), we have

$$\frac{y_A}{x_A} = \frac{w_y}{\bar{x}}. \quad (11.26)$$

---

<sup>2</sup>As we discussed above, this is true if the consumption externality is positive, or there is no externality or only one side externality.

Then, by (11.25) and (11.26), we have

$$\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} = \frac{1}{4} \sqrt{\frac{y_A}{x_A}} \sqrt{\frac{y_B}{x_B}} = \frac{y_A}{4x_A} = \frac{w_y}{4\bar{x}} \quad (11.27)$$

and

$$\frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A} = 1. \quad (11.28)$$

Thus,  $\bar{x} = w_y/4$  is the critical point that makes  $\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} - \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A} = 0$ , or equivalently  $\frac{\frac{\partial u_A}{\partial x_A}}{\frac{\partial u_A}{\partial y_A}} + \frac{\frac{\partial u_B}{\partial x_A}}{\frac{\partial u_B}{\partial y_B}} = \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} + \frac{\frac{\partial u_A}{\partial x_B}}{\frac{\partial u_A}{\partial y_A}} = 0$ . Hence, if  $w_x > \frac{w_y}{4}$ , then  $\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} - \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A} < 0$ , and thus, by (11.22), there is destruction in any Pareto efficient allocation. If  $w_x < \frac{w_y}{4}$ , then  $\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} - \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A} > 0$ , and, by (11.23), no Pareto optimal allocation requires destruction. Finally, when  $w_x = \frac{w_y}{4}$ , any allocation that satisfies the marginal equality condition (11.15) and the balanced conditions  $x_A + x_B = w_x$  and  $y_A + y_B = w_y$  also satisfies (11.19), and thus it is a Pareto efficient allocation with no destruction.

Now, let us set out the sufficient conditions for destruction and non-destruction in detail. Note that if the sufficient condition for non-destruction holds, then  $x_A + x_B = w_x$  would also hold as an implication. Thus, to apply both the sufficient conditions for non-destruction and destruction, one should use the following three conditions

$$\left\{ \begin{array}{l} \frac{\frac{\partial u_A}{\partial x_A}}{\frac{\partial u_A}{\partial y_A}} + \frac{\frac{\partial u_B}{\partial x_A}}{\frac{\partial u_B}{\partial y_B}} = \frac{\frac{\partial u_B}{\partial x_B}}{\frac{\partial u_B}{\partial y_B}} + \frac{\frac{\partial u_A}{\partial x_B}}{\frac{\partial u_A}{\partial y_A}} \\ x_A + x_B = w_x \\ y_A + y_B = w_y \end{array} \right. .$$

If an allocation also satisfies the condition  $\frac{\partial u_A}{\partial x_A} \frac{\partial u_B}{\partial x_B} \geq \frac{\partial u_A}{\partial x_B} \frac{\partial u_B}{\partial x_A}$ , then the allocation is Pareto efficient, and further there is no destruction. Otherwise, it is not Pareto efficient. If this is true for all such non-Pareto allocations, all Pareto efficient condition must hold with destruction.

Note that, since  $\frac{\partial u_A}{\partial x_A}$  and  $\frac{\partial u_B}{\partial x_B}$  represent marginal benefit, they are usually diminishing in consumption in good  $x$ . Since  $\frac{\partial u_A}{\partial x_B}$  and  $\frac{\partial u_B}{\partial x_A}$  are in the form of a marginal cost, their absolute values would be typically increasing in the consumption of good  $x$ . Hence, when total endowment  $w_x$  is small, the social marginal benefit would exceed the social marginal cost so that there is no destruction. As the total endowment of  $w_x$  increases, the social marginal cost will ultimately outweigh the marginal social benefit, which results in the destruction of the endowment of  $w_x$ .

Alternatively, we can get the same result by using social marginal rates of substitution. When utility functions are strictly quasi-concave, marginal rates of substitution are diminishing. Therefore, in the presence of negative consumption externalities, social marginal rates of substitution may become negative when the consumption of good  $x$  becomes sufficiently large. When this occurs, it is better to destroy some resources for good  $x$ . As the destruction of good  $x$  increases, which will, in turn, decrease the consumption of good  $x$ , social marginal rates of substitution will increase. Eventually they will become nonnegative.

The destruction issue is not only important in theory, but also relevant to reality. It can be used to explain a well-known puzzle of the happiness-income relationship in the economic and psychology literatures: happiness rises with income up to a point, but not beyond it. For example, well-being has declined over the last quarter of a century in the U.S., and life satisfaction has run approximately flat across the same time in Britain. If we interpret income as a good, and if consumers envy each other in terms of consumption levels, by our result, when the income reaches a certain level, one may have to freely dispose of wealth to achieve Pareto efficient allocations; otherwise the resulting allocations will be Pareto inefficient. Therefore, economic growth does not raise well-being indexed by any social welfare function once the critical income level is achieved. For detailed discussion, see Tian and Yang (2007, 2009).

### 11.3 Production Externality

We now show that allocation of resources may not be efficient also for the case of externality in production. To show this, consider a simple economy with two firms. Firm 1 produces an output  $x$  which will be sold in a competitive market. However, production of  $x$  imposes an externality cost denoted by  $e(x)$  to firm 2, which is assumed to be convex and strictly increasing.

Let  $y$  be the output produced by firm 2, which is sold in competitive market.

Let  $c_x(x)$  and  $c_y(y)$  be the cost functions of firms 1 and 2 which are both convex and strictly increasing.

The profits of the two firms:

$$\pi_1 = p_x x - c_x(x) \quad (11.29)$$

$$\pi_2 = p_y y - c_y(y) - e(x) \quad (11.30)$$

where  $p_x$  and  $p_y$  are the prices of  $x$  and  $y$ , respectively. Then, by the first order conditions, we have for positive amounts of outputs:

$$p_x = c'_x(x) \quad (11.31)$$

$$p_y = c'_y(y) \quad (11.32)$$

However, the profit maximizing output  $x_c$  from the first order condition is too large from a social point of view. The first firm only takes account of the private cost – the cost that is imposed on itself– but it ignores the social cost – the private cost plus the cost it imposes on the other firm.

What's the social efficient output?

The social profit,  $\pi_1 + \pi_2$ , is not maximized at  $x_c$  and  $y_c$  which satisfy (11.31) and (11.32). If the two firms merged so as to internalize the externality

$$\max_{x,y} p_x x + p_y y - c_x(x) - e(x) - c_y(y) \quad (11.33)$$

which gives the first order conditions:

$$p_x = c'_x(x^*) + e'(x^*)$$

$$p_y = c'_y(y^*)$$

where  $x^*$  is an efficient amount of output; it is characterized by price being equal to the marginal social cost. Thus, production of  $x^*$  is less than the competitive output in the externality case by the convexity of  $e(x)$  and  $c_x(x)$ .



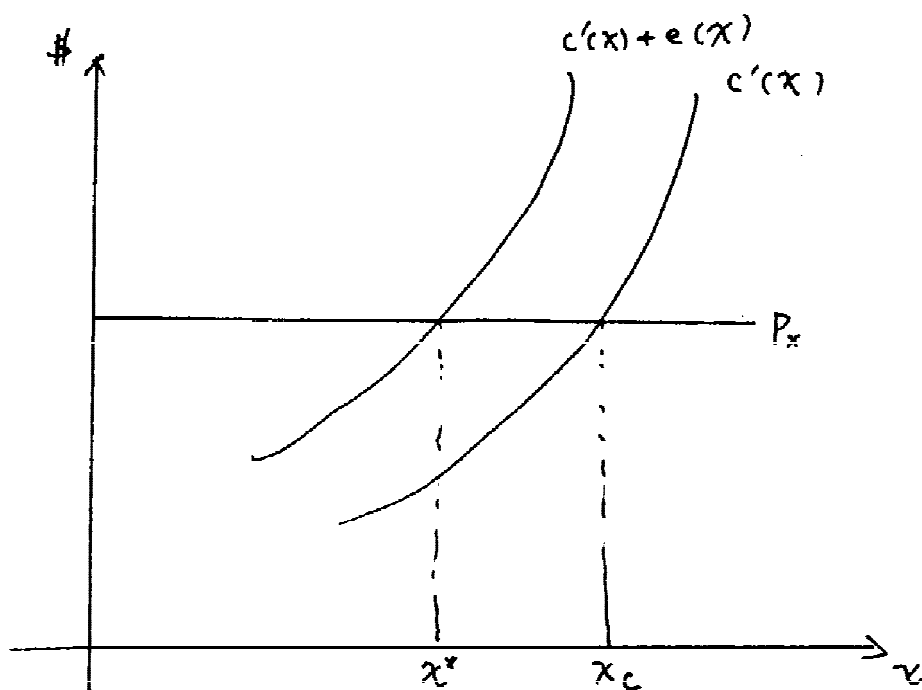


Figure 11.1: The efficient output  $x^*$  is less than the competitive output  $x_c$ .

## 11.4 Solutions to Externalities

From the above discussion, we know that a competitive market in general may not result in Pareto efficient outcome in the presence of externalities, and one needs to seek some other alternative mechanisms to solve the market failure problem. In this section, we now introduce some remedies to this market failure of externality such as:

1. Pigovian taxes
2. Voluntary negotiation (Coase Approach)
3. Compensatory tax/subsidy
4. Creating a missing markets with property rights
5. Direct intervention
6. Merges
7. Incentives mechanism design

Any of the above solution may result in Pareto efficient outcomes, but may lead to different income distributions. Also, it is important to know what kind of information are required to implement a solution listed above.

Most of the above proposed solutions need to make the following assumptions:

1. The source and degree of the externality is identifiable.
2. The recipients of the externality is identifiable.
3. The causal relationship of the externality can be established objectively.
4. The cost of preventing (by different methods) an externality are perfectly known to everyone.
5. The cost of implementing taxes and subsidies is negligible.
6. The cost of voluntary negotiation is negligible.

### 11.4.1 Pigovian Tax

Set a tax rate,  $t$ , such that  $t = e'(x^*)$ . This tax rate to firm 1 would internalize the externality.

$$\pi_1 = p_x \cdot x - c_x(x) - t \cdot x \quad (11.34)$$

The first order condition is:

$$p_x = c'_x(x) + t = c'_x(x) + e'(x^*), \quad (11.35)$$

which is the same as the one for social optimality. That is, when firm 1 faces the wrong price of its action, and a tax  $t = e'(x^*)$  should be imposed for each unit of firm 1's production. This will lead to a social optimal outcome that is less than that of competitive equilibrium outcome. Such correction taxes are called Pigovian taxes.

The problem with this solution is that it requires that the taxing authority knows the externality cost  $e(x)$ . But, how does the authority know the externality and how do they estimate the value of externality in real world? If the authority knows this information, it might as well just tell the firm how much to produce in the first place. So, in most case, it will not work well.

## 11.4.2 Coase' Voluntary Negotiation and Enforceable Property Rights Approach

A different approach to the externality problem relies on the parties to negotiate a solution to the problem themselves.

The greatest novelty of Nobel laureate Ronald Coase's contribution was the systematic treatment of trade in property rights. To solve the externality problem, Coase in a famous article, "The Problem of Social Cost", in 1960 argues that the success of such a system depends on making sure that property rights are clearly assigned. The so-called *Coases Theorem* assesses that as long as property rights are clearly assigned, the two parties will negotiate in such a way that the optimal level of the externality-producing activity is implemented. As a policy implication, a government should simply rearrange property rights with appropriately designed property rights. Market then could take care of externalities without direct government intervention.

Coase theorem thus contains two claims. One is that the level of the externality will be the same regardless of the assignment of property rights, which is called *Coase Neutrality Theorem*. The second one is that voluntary negotiations over externalities will lead to a Pareto-optimal outcome, which is called *Coase Efficiency Theorem*.

Coase shows his propositions mainly by using examples of economy with only two agents and detrimental externality. Coase made an observation that in the presence of externalities, the victim has an incentive to pay the firm to stop the production if the victim can compensate the firm by paying  $p_x - c'_x(x^*)$ . The following example captures Coase's arguments.

**Example 11.4.1** Two firms: One is chemical factory that discharges chemicals into a river and the other is the fisherman. Suppose the river can produce a value of \$50,000. If the chemicals pollute the river, the fish cannot be eaten. How does one solve the externality? Coase's method states that as long as the property rights of the river are clearly assigned, it results in efficient outcomes. That is, the government should give the ownership of the lake either to the chemical firm or to the fisherman, then it will yield an efficient output. To see this, assume that:

The cost of the filter is denoted by  $c_f$ .

*Case 1:* The lake is given to the factory.

- i)  $c_f < \$50,000$ . The fisherman is willing to buy a filter for the factory. The fisherman will pay for the filter so that the chemical cannot pollute the lake.
- ii)  $c_f > \$50,000$  – The chemical is discharged into the lake. The fisherman does not want to install any filter.

*Case 2:* The lake is given to the fisherman, and the firm's net product revenue is greater than \$50,000.

- i)  $c_f < \$50,000$  – The factory buys the filter so that the chemical cannot pollute the lake.
- ii)  $c_f > \$50,000$  – The firm pays \$50,000 to the fisherman then the chemical is discharged into the lake.

Like the above example, Coase's own examples supporting his claims do with negotiations between firms or business rather than those between individuals. This difference is important since firms maximize profits rather than utility, and act as fiduciaries for individuals. We have to make some restricted assumptions on consumers' utility functions to make Coase's Theorem to be held.

Now consider an economy with two consumers with  $L$  goods. Further, consumer  $i$  has initial wealth  $w_i$ . Each consumer has preferences over both the commodities he consumes and over some action  $h$  that is taken by consumer 1. That is,

$$u_i(x_i^1, \dots, x_i^L, h).$$

Activity  $h$  is something that has no direct monetary cost for person 1. For example,  $h$  is the quantity of loud music played by person 1. In order to play it, the consumer must purchase electricity, but electricity can be captured as one of the components of  $x_i$ . From the point of view of consumer 2,  $h$  represents an external effect of consumer 1's action. In the model, we assume that

$$\frac{\partial u_2}{\partial h} \neq 0.$$

Thus the externality in this model lies in the fact that  $h$  affects consumer 2's utility,

but it is not priced by the market. Let  $v_i(p, w_i, h)$  be consumer  $i$ 's indirect utility function:

$$\begin{aligned} v_i(w_i, h) &= \max_{x_i} u_i(x_i, h) \\ \text{s.t.} \quad &px_i \leq w_i. \end{aligned}$$

We assume that preferences are quasi-linear with respect to some numeraire commodity. Thus, the consumer's indirect utility function takes the form:

$$v_i(w_i, h) = \phi_i(h) + w_i.$$

We further assume that utility is strictly concave in  $h$  :  $\phi_i''(h) < 0$ . Again, the competitive equilibrium outcome in general is not Pareto optimal. In order to maximize utility, the consumer 1 should choose  $h$  in order to maximize  $v_1$  so that the interior solution satisfies  $\phi_1'(h^*) = 0$ . Even though consumer 2's utility depends on  $h$ , it cannot affect the choice of  $h$ .

On the other hand, the socially optimal level of  $h$  will maximize the sum of the consumers' utilities:

$$\max_h \phi_1(h) + \phi_2(h).$$

The first-order condition for an interior maximum is:

$$\phi_1'(h^{**}) + \phi_2'(h^{**}) = 0,$$

where  $h^{**}$  is the Pareto optimal amount of  $h$ . Thus, the social optimum is where the sum of the marginal benefit of the two consumers equals zero. In the case where the externality is bad for consumer 2 (loud music), the level of  $h^* > h^{**}$ . That is, too much  $h$  is produced. In the case where the externality is good for consumer 2 (baking bread smell or yard beautification), too little will be provided,  $h^* < h^{**}$ .

Now we show that, as long as property rights are clearly assigned, the two parties will negotiate in such a way that the optimal level of the externality-producing activity is implemented. We first consider the case where consumer 2 has the right to prohibit consumer 1 from undertaking activity  $h$ . But, this right is contractible. Consumer 2 can sell consumer 1 the right to undertake  $h_2$  units of activity  $h$  in exchange for some transfer,

$T_2$ . The two consumers will bargain both over the size of the transfer  $T_2$  and over the number of units of the externality good produced,  $h_2$ .

In order to determine the outcome of the bargaining, we first specify the bargaining mechanism as follows:

1. Consumer 2 offers consumer 1 a take-it-or-leave-it contract specifying a payment  $T_2$  and an activity level  $h_2$ .
2. If consumer 1 accepts the offer, that outcome is implemented. If consumer 1 does not accept the offer, consumer 1 cannot produce any of the externality good, i.e.,  $h = 0$ .

To analyze this, begin by considering which offers  $(h, T)$  will be accepted by consumer 1. Since in the absence of agreement, consumer 1 must produce  $h = 0$ , consumer 1 will accept  $(h_2, T_2)$  if and only if it offers higher utility than  $h = 0$ . That is, consumer 1 accepts if and only if:

$$\phi_1(h_2) - T_2 \geq \phi_1(0).$$

Given this constraint on the set of acceptable offers, consumer 2 will choose  $(h_2, T_2)$  in order to solve the following problem:

$$\begin{aligned} \max_{h_2, T_2} & \phi_2(h_2) + T_2 \\ \text{s.t.} & : \phi_1(h_2) - T_2 \geq \phi_1(0). \end{aligned}$$

Since consumer 2 prefers higher  $T_2$ , the constraint will bind at the optimum. Thus the problem becomes:

$$\max_{h_2} \phi_1(h_2) + \phi_2(h_2) - \phi_1(0).$$

The first-order condition for this problem is given by:

$$\phi'_1(h_2) + \phi'_2(h_2) = 0.$$

But, this is the same condition that defines the socially optimal level of  $h_2$ . Thus consumer 2 chooses  $h_2 = h^{**}$ , and, using the constraint,  $T_2 = \phi_1(h^{**}) - \phi_1(0)$ . And, the offer  $(h_2, T_2)$  is accepted by consumer 1. Thus this bargaining process implements the social optimum.

Now we consider the case where consumer 1 has the right to produce as much of the externality as she wants. We maintain the same bargaining mechanism. Consumer 2 makes consumer 1 a take-it-or-leave-it offer  $(h_1, T_1)$ , where the subscript indicates that consumer 1 has the property right in this situation. However, now, in the event that 1 rejects the offer, consumer 1 can choose to produce as much of the externality as she wants, which means that she will choose to produce  $h^*$ . Thus the only change between this situation and the first case is what happens in the event that no agreement is reached. In this case, consumer 2's problem is:

$$\begin{aligned} \max_{h_1, T_1} & \phi_2(h_1) - T_1 \\ \text{s.t.} & : \phi_1(h_1) + T_1 \geq \phi_1(h^*) \end{aligned}$$

Again, we know that the constraint will bind, and so consumer 2 chooses  $h_1$  and  $T_1$  in order to maximize

$$\max \phi_1(h_1) + \phi_2(h_1) - \phi_1(h^*)$$

which is also maximized at  $h_1 = h^{**}$ , since the first-order condition is the same. The only difference is in the transfer. Here  $T_1 = \phi_1(h^*) - \phi_1(h^{**})$ .

While both property-rights allocations implement  $h^{**}$ , they have different distributional consequences. The payment of the transfer is positive in the case where consumer 2 has the property rights, while it is negative when consumer 1 has the property rights. The reason for this is that consumer 2 is in a better bargaining position when the non-bargaining outcome is that consumer 1 is forced to produce 0 units of the externality good.

However, note that in the quasilinear framework, redistribution of the numeraire commodity has no effect on social welfare. The fact that regardless of how the property rights are allocated, bargaining leads to a Pareto optimal allocation is an example of the Coase Theorem: If trade of the externality can occur, then bargaining will lead to an efficient outcome no matter how property rights are allocated (as long as they are clearly allocated). Note that well-defined, enforceable property rights are essential for bargaining to work. If there is a dispute over who has the right to pollute (or not pollute), then bargaining may not lead to efficiency. An additional requirement for efficiency is that the bargaining process itself is costless. Note that the government doesn't need to know about

individual consumers here - it only needs to define property rights. However, it is critical that it do so clearly. Thus the Coase Theorem provides an argument in favor of having clear laws and well-developed courts.

However, Hurwicz (Japan and the World Economy 7, 1995, pp. 49-74) argued that, even when the transaction cost is zero, absence of income effects in the demand for the good with externality is not only sufficient (which is well known) but also necessary for Coase Neutrality Theorem to be true, i.e., when the transaction cost is negligible, the level of pollution will be independent of the assignments of property rights if and only if preferences of the consumers are quasi-linear with respect to the private good, leading to absence of income effects in the demand for the good with externality.

Unfortunately, as shown by Chipman and Tian (2012, Economic Theory), the proof of Hurwicz's claim on the necessity of parallel preferences for "Coase's conjecture" is incorrect. To see this, consider the following class of utility functions  $U_i(x_i, h)$  that have the functional form:

$$U_i(x_i, h) = x_i e^{-h} + \phi_i(h), \quad i = 1, 2 \quad (11.36)$$

where

$$\phi_i(h) = \int e^{-h} b_i(h) dh. \quad (11.37)$$

$U_i(x_i, h)$  is then clearly not quasi-linear in  $x_i$ . It is further assumed that for all  $h \in (0, \eta]$ ,  $b_1(h) > \xi$ ,  $b_2(h) < 0$ ,  $b'_i(h) < 0$  ( $i = 1, 2$ ),  $b_1(0) + b_2(0) \geq \xi$ , and  $b_1(\eta) + b_2(\eta) \leq \xi$ .

We then have

$$\begin{aligned} \partial U_i / \partial x_i &= e^{-h} > 0, \quad i = 1, 2, \\ \partial U_1 / \partial h &= -x_1 e^{-h} + b_1(h) e^{-h} > e^{-h} [\xi - x_1] \geq 0, \\ \partial U_2 / \partial h &= -x_2 e^{-h} + b_2(h) e^{-h} < 0 \end{aligned}$$

for  $(x_i, h) \in (0, \xi) \times (0, \eta)$ ,  $i = 1, 2$ . Thus, by the mutual tangency equality condition for Pareto efficiency, we have

$$0 = \frac{\partial U_1}{\partial h} \bigg/ \frac{\partial U_1}{\partial x_1} + \frac{\partial U_2}{\partial h} \bigg/ \frac{\partial U_2}{\partial x_2} = -x_1 - x_2 + b_1(h) + b_2(h) = b_1(h) + b_2(h) - \xi, \quad (11.38)$$

which is independent of  $x_i$ . Hence, if  $(x_1, x_2, h)$  is Pareto optimal, so is  $(x'_1, x'_2, h)$  provided  $x_1 + x_2 = x'_1 + x'_2 = \xi$ . Also, note that  $b'_i(h) < 0$  ( $i = 1, 2$ ),  $b_1(0) + b_2(0) \geq \xi$ , and  $b_1(\eta) + b_2(\eta) \leq \xi$ . Then  $b_1(h) + b_2(h)$  is strictly monotone and thus there is a unique  $h \in$



$[0, \eta]$ , satisfying (11.38). Thus, the contract curve is horizontal even though individuals' preferences need not be parallel.

**Example 11.4.2** Suppose  $b_1(h) = (1+h)^\alpha \eta^\eta + \xi$  with  $\alpha < 0$ , and  $b_2(h) = -h^\eta$ . Then, for all  $h \in (0, \eta]$ ,  $b_1(h) > \xi$ ,  $b_2(h) < 0$ ,  $b'_i(h) < 0$  ( $i = 1, 2$ ),  $b_1(0) + b_2(0) > \xi$ , and  $b_1(\eta) + b_2(\eta) < \xi$ . Thus,  $\phi_i(h) = \int e^{-h} b_i(h) dh$  is concave, and  $U_i(x_i, h) = x_i e^{-h} + \int e^{-h} b_i(h) dh$  is quasi-concave,  $\partial U_i / \partial x_i > 0$  and  $\partial U_1 / \partial h > 0$ , and  $\partial U_2 / \partial h < 0$  for  $(x_i, h) \in (0, \xi) \times (0, \eta)$ ,  $i = 1, 2$ , but it is not quasi-linear in  $x_i$ .

Chipman and Tian (2012) then investigate the necessity for the ‘‘Coase conjecture’’ that the level of pollution is independent of the assignments of property rights. This reduces to developing the necessary and sufficient conditions that guarantee that the contract curve is horizontal so that the set of Pareto optima for the utility functions is  $h$ -constant. This in turn reduces to finding the class of utility functions such that the mutual tangency (first-order) condition does not contain  $x_i$  and consequently it is a function, denoted by  $g(h)$ , of  $h$  only:

$$\frac{\partial U_1}{\partial h} \bigg/ \frac{\partial U_1}{\partial x_1} + \frac{\partial U_2}{\partial h} \bigg/ \frac{\partial U_2}{\partial x_2} = g(h) = 0. \quad (11.39)$$

Let  $F_i(x_i, h) = \frac{\partial U_i}{\partial h} / \frac{\partial U_i}{\partial x_i}$  ( $i = 1, 2$ ), which can be generally expressed as

$$F_i(x_i, h) = x_i \psi_i(h) + f_i(x_i, h) + b_i(h),$$

where  $f_i(x_i, h)$  are nonseparable and nonlinear in  $x_i$ .  $\psi_i(h)$ ,  $b_i(h)$ , and  $f_i(x_i, h)$  will be further specified below.

Let  $F(x, h) = F_1(x, h) + F_2(\xi - x, h)$ . Then the mutual tangency equality condition can be rewritten as

$$F(x, h) = 0. \quad (11.40)$$

Thus, the contract curve, i.e., the locus of Pareto-optimal allocations, can be expressed by a function  $h = f(x)$  that is implicitly defined by (11.40).

Then, the Coase Neutrality Theorem, which is characterized by the condition that the set of Pareto optima (the contract curve) in the  $(x, h)$  space for  $x_i > 0$  is a horizontal line  $h = \text{constant}$ , implies that

$$h = f(x) = \bar{h}$$

with  $\bar{h}$  constant, and thus we have

$$\frac{dh}{dx} = -\frac{F_x}{F_h} = 0$$

for all  $x \in [0, \xi]$  and  $F_h \neq 0$ , which means that the function  $F(x, h)$  is independent of  $x$ . Then, for all  $x \in [0, \xi]$ ,

$$F(x, h) = x\psi_1(h) + (\xi - x)\psi_2(h) + f_1(x, h) + f_2(\xi - x, h) + b_1(h) + b_2(h) \equiv g(h). \quad (11.41)$$

Since the utility functions  $U_1$  and  $U_2$  are functionally independent, and  $x$  disappears in (11.41), we must have  $\psi_1(h) = \psi_2(h) \equiv \psi(h)$  and  $f_1(x, h) = -f_2(\xi - x, h) = 0$  for all  $x \in [0, \xi]$ . Therefore,

$$F(x, h) = \xi\psi(h) + b_1(h) + b_2(h) \equiv g(h), \quad (11.42)$$

and

$$\frac{\partial U_i}{\partial h} \bigg/ \frac{\partial U_i}{\partial x_i} = F_i(x_i, h) = x_i\psi(h) + b_i(h) \quad (11.43)$$

which is a first-order linear partial differential equation. Then, from Polyanin, Zaitsev, and Moussiaux (2002),<sup>3</sup> we know that the principal integral  $U_i(x_i, h)$  of (11.43) is given by

$$U_i(x_i, h) = x_i e^{\int \psi(h)} + \phi_i(h), \quad i = 1, 2 \quad (11.44)$$

with

$$\phi_i(h) = \int e^{\int \psi(h)} b_i(h) dh. \quad (11.45)$$

The general solution of (11.43) is then given by  $\bar{U}_i(x, y) = \psi(U_i)$ , where  $\psi$  is an arbitrary function. Since a monotonic transformation preserves orderings of preferences, we can regard the principal solution  $U_i(x_i, h)$  as a general functional form of utility functions that is fully characterized by (11.43).

Note that (11.44) is a general utility function that contains quasi-linear utility in  $x_i$  and the utility function given in (11.36) as special cases. Indeed, it represents parallel preferences when  $\psi(h) \equiv 0$  and also reduces to the utility function given by (11.36) when  $\psi(h) = -1$ .

To make the the mutual tangency (first-order) condition (11.39) be also sufficient for the contract curve to be horizontal in a pollution economy, we assume that for all

---

<sup>3</sup>It can be also seen from <http://eqworld.ipmnet.ru/en/solutions/fpde/fpde1104.pdf>.

$h \in (0, \eta]$ ,  $x_1\psi(h) + b_1(h) > 0$ ,  $x_2\psi(h) + b_2(h) < 0$ ,  $\psi'(h) \leq 0$ ,  $b'_i(h) < 0$  ( $i = 1, 2$ ),  $\xi\psi(0) + b_1(0) + b_2(0) \geq 0$ , and  $\xi\psi(\eta) + b_1(\eta) + b_2(\eta) \leq 0$ .

We then have for  $(x_i, h) \in (0, \xi) \times (0, \eta)$ ,  $i = 1, 2$ ,

$$\begin{aligned}\partial U_i / \partial x_i &= e^{\int \psi(h)} > 0, \quad i = 1, 2, \\ \partial U_1 / \partial h &= e^{\int \psi(h)} [x_1\psi(h) + b_1(h)] > 0, \\ \partial U_2 / \partial h &= e^{\int \psi(h)} [x_2\psi(h) + b_2(h)] < 0,\end{aligned}$$

and thus

$$0 = \frac{\partial U_1}{\partial h} \bigg/ \frac{\partial U_1}{\partial x_1} + \frac{\partial U_2}{\partial h} \bigg/ \frac{\partial U_2}{\partial x_2} = (x_1 + x_2)\psi(h) + b_1(h) + b_2(h) = \xi\psi(h) + b_1(h) + b_2(h), \quad (11.46)$$

which does not contain  $x_i$ . Hence, if  $(x_1, x_2, h)$  is Pareto optimal, so is  $(x'_1, x'_2, h)$  provided  $x_1 + x_2 = x'_1 + x'_2 = \xi$ . Also, note that  $\psi'(h) \leq 0$ ,  $b'_i(h) < 0$  ( $i = 1, 2$ ),  $\xi\psi(0) + b_1(0) + b_2(0) \geq 0$ , and  $\xi\psi(\eta) + b_1(\eta) + b_2(\eta) \leq 0$ . Then  $\xi\psi(h) + b_1(h) + b_2(h)$  is strictly monotone and thus there is a unique  $h \in [0, \eta]$  that satisfies (11.46). Thus, the contract curve is horizontal even though individuals' preferences need not be parallel.

The formal statement of Coase Neutrality Theorem obtained by Chipman and Tian (2012) thus can be set forth as follows:

**Proposition 11.4.1 (Coase Neutrality Theorem)** *In a pollution economy considered in the chapter, suppose that the transaction cost equals zero, and that the utility functions  $U_i(x_i, h)$  are differentiable and such that  $\partial U_i / \partial x_i > 0$ , and  $\partial U_1 / \partial h > 0$  but  $\partial U_2 / \partial h < 0$  for  $(x_i, h) \in (0, \xi) \times (0, \eta)$ ,  $i = 1, 2$ . Then, the level of pollution is independent of the assignments of property rights if and only if the utility functions  $U_i(x, y)$ , up to a monotonic transformation, have a functional form given by*

$$U_i(x_i, h) = x_i e^{\int \psi(h)} + \int e^{\int \psi(h)} b_i(h) dh, \quad (11.47)$$

where  $h$  and  $b_i$  are arbitrary functions such that the  $U_i(x_i, h)$  are differentiable,  $\partial U_i / \partial x_i > 0$ , and  $\partial U_1 / \partial h > 0$  but  $\partial U_2 / \partial h < 0$  for  $(x_i, h) \in (0, \xi) \times (0, \eta)$ ,  $i = 1, 2$ .

Although the above Coase neutrality theorem covers a wider class of preferences than quasi-linear utility functions, it still puts a significant restriction on the domain of its validity due to the special functional forms of the utility functions with respect to the

private good. Thus, the Coase Neutrality Theorem may be best applied to production externalities between firms, but rather than consumption externalities.

The problem of this Coase theorem is that, costs of negotiation and organization, in general, are not negligible, and the income effect may not be zero. Thus, a privatization is optimal only in case of zero transaction cost, no income effect, and perfect economic environments. Tian (2000, 2001) shown that the state ownership and the collective ownership can also be optimal when an economic environment is sufficient imperfect or somewhere in the middle.

The problem of the Coase Efficiency Theorem is more serious. First, as Arrow (1979, p. 24) pointed out, the basic postulate underlying Coase's theory appears to be that the process of negotiation over property rights can be modelled as a cooperative game, and this requires the assumption that each player know the preferences or production functions of each of the other players. When information is not complete or asymmetric, in general we do not have a Pareto optimal outcome. For instance, when there is one polluter and there are many pollutees, a "free-rider" problem arises and there is an incentive for pollutees to misrepresent their preferences. Whether the polluter is liable or not, the pollutees may be expected overstate the amount they require to compensate for the externality. Thus, we may need to design an incentive mechanism to solve the free-rider problem.

Secondly, even if the information is complete, there are several circumstances that have led a number of authors to question the conclusion in the Coase Efficiency Theorem:

- (1) The core may be empty, and hence no Pareto optimum exists. An example of this for a three-agent model was presented by Aivazian and Callen (1981).
- (2) There may be a fundamental non-convexity that prevents a Pareto optimum from being supported by a competitive equilibrium. Starrett (1972) showed that externalities are characterized by "fundamental non-convexities" that may preclude existence of competitive equilibrium.
- (3) When an agent possesses the right to pollute, there is a built-in incentive for extortion. As Andel (1966) has pointed out, anyone with the right to pollute has an incentive to extract payment from potential pollutees, e.g.,

threat to blow a bugle in the middle of the night.

Thus, the hypothesis that negotiations over externalities will mimic trades in a competitive equilibrium is, as Coase himself has conceded, not one that can be logically derived from his assumptions, but must be regarded as an empirical conjecture that may or may not be confirmed from the data. A lot of theoretical work therefore still remains in order to provide Coasian economics with the rigorous underpinning.

### 11.4.3 Missing Market

We can regard externality as a lack of a market for an “externality.” For the above example in Pigovian taxes, a missing market is a market for pollution. Adding a market for firm 2 to express its demand for pollution - or for a reduction of pollution - will provide a mechanism for efficient allocations. By adding this market, firm 1 can decide how much pollution it wants to sell, and firm 2 can decide how much pollution it wants to buy.

Let  $r$  be the price of pollution.

$x_1$  = the units of pollution that firm 1 wants to sell;

$x_2$  = the units of pollution for firm 2 wants to buy.

Normalize the output of firm 1 to  $x_1$ .

The profit maximization problems become:

$$\begin{aligned}\pi_1 &= p_x x_1 + r x_1 - c_1(x_1) \\ \pi_2 &= p_y y - r x_2 - e_2(x_2) - c_y(y)\end{aligned}$$

The first order conditions are:

$$\begin{aligned}p_x + r &= c'_1(x_1) \quad \text{for Firm 1} \\ p_y &= c'_y(y) \quad \text{for Firm 2} \\ -r &= e'(x_2) \quad \text{for Firm 2.}\end{aligned}$$

At the market equilibrium,  $x_1^* = x_2^* = x^*$ , we have

$$p_x = c'_1(x^*) + e'(x^*) \tag{11.48}$$

which results in a social optimal outcome.

### 11.4.4 The Compensation Mechanism

The Pigovian taxes were not adequate in general to solve externalities due to the information problem: the tax authority cannot know the cost imposed by the externality. How can one solve this incomplete information problem?

Varian (AER 1994) proposed an incentive mechanism which encourages the firms to correctly reveal the costs they impose on the other. Here, we discuss this mechanism. In brief, a mechanism consists of a message space and an outcome function (rules of game). We will introduce in detail the mechanism design theory in Part V.

**Strategy Space (Message Space):**  $M = M_1 \times M_2$  with  $M_1 = \{(t_1, x_1)\}$ , where  $t_1$  is interpreted as a Pigovian tax proposed by firm 1 and  $x_1$  is the proposed level of output by firm 1, and  $t_2$  is interpreted as a Pigovian tax proposed by firm 2 and  $y_2$  is the proposed level of output by firm 2.

The mechanism has two stages:

*Stage 1:* (Announcement stage): Firms 1 and 2 name Pigovian tax rates,  $t_i$ ,  $i = 1, 2$ , which may or may not be the efficient level of such a tax rate.

*Stage 2:* (Choice stage): If firm 1 produces  $x$  units of pollution, firm 1 must pay  $t_2x$  to firm 2. Thus, each firm takes the tax rate as given. Firm 2 receives  $t_1x$  units as compensation. Each firm pays a penalty,  $(t_1 - t_2)^2$ , if they announce different tax rates.

Thus, the payoffs of two firms are:

$$\begin{aligned}\pi_1^* &= \max_x p_x x - c_x(x) - t_2 x - (t_1 - t_2)^2 \\ \pi_2^* &= \max_y p_y y - c_y(y) + t_1 x - e(x) - (t_1 - t_2)^2.\end{aligned}$$

Because this is a two-stage game, we may use the subgame perfect equilibrium, i.e., an equilibrium in which each firm takes into account the repercussions of its first-stage choice on the outcomes in the second stage. As usual, we solve this game by looking at stage 2 first.

At stage 2, firm 1 will choose  $x(t_2)$  to satisfy the first order condition:

$$p_x - c'_x(x) - t_2 = 0 \quad (11.49)$$

Note that, by the convexity of  $c_x$ , i.e.,  $c''_x(x) > 0$ , we have

$$x'(t_2) = -\frac{1}{c''_x(x)} < 0. \quad (11.50)$$

Firm 2 will choose  $y$  to satisfy  $p_y = c'_y(y)$ .

*Stage 1:* Each firm will choose the tax rate  $t_1$  and  $t_2$  to maximize their payoffs.

For Firm 1,

$$\max_{t_1} p_x x - c_x(x) - t_2 x(t_2) - (t_1 - t_2)^2 \quad (11.51)$$

which gives us the first order condition:

$$2(t_1 - t_2) = 0$$

so the optimal solution is

$$t_1^* = t_2. \quad (11.52)$$

For Firm 2,

$$\max_{t_2} p_y y - c_y(y) + t_1 x(t_2) - e(x(t_2)) - (t_1 - t_2)^2 \quad (11.53)$$

so that the first order condition is

$$t_1 x'(t_2) - e'(x(t_2))x'(t_2) + 2(t_1 - t_2) = 0$$

and thus

$$[t_1 - e'(x(t_2))]x'(t_2) + 2(t_1 - t_2) = 0. \quad (11.54)$$

By (11.50), (11.52) and (11.54), we have

$$t^* = e'(x(t^*)) \quad \text{with} \quad t^* = t_1^* = t_2^*. \quad (11.55)$$

Substituting the equilibrium tax rate,  $t^* = e'(x(t^*))$ , into (11.49) we have

$$p_x = c'_x(x^*) + e'(x^*) \quad (11.56)$$

which is the condition for social efficiency of production.

**Remark 11.4.1** This mechanism works by setting opposing incentives for two agents. Firm 1 always has an incentive to match the announcement of firm 2. But consider firm 2's incentive. If firm 2 thinks that firm 1 will propose a large compensation rate  $t_1$  for him, he wants firm 1 to be taxed as little as possible so that firm 1 will produce as much as possible. On the other hand, if firm 2 thinks firm 1 will propose a small  $t_1$ , it wants firm 1 to be taxed as much as possible. Thus, the only point where firm 2 is indifferent about the level of production of firm 1 is where firm 2 is exactly compensated for the cost of the externality.

In general, the individual's objective is different from the social goal. However, we may be able to construct an appropriated mechanism so that the individual's profit maximizing goal is consistent with the social goal such as efficient allocations. Tian (2003) also gave the solution to the consumption externalities by giving the incentive mechanism that results in Pareto efficient allocations. Tian (2004) study the informational efficiency problem of the mechanisms that results in Pareto efficient allocations for consumption externalities.

## Reference

- Andel, N., "Some Note on Equating Private and Social Cost: Comment," *Southern Economic Journal*, 33 (1966), 112.
- Aivazian, V. A., and J. L. Callen, "The Coase Theorem and the Empty Core," *Journal of Law and Economics*, 24 (1981), 175-181.
- Arrow, K. J., "The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Nonmuarket Allocation," in: *The Analysis and Evaluation of public Expenditures: The PPB System*, 1969. 3.
- Arrow, K. J., "The Property Rights Doctrine and Demand Revelation under Incomplete Information," in Michael J. Boskin (ed.), *Economics and Human Welfare: Essays in Honor Tibor Scitovsky*, New York: Academic Press, 23-39.
- Chipman, J. S., "A Close Look to the Coase Theorem," in *The Economists' Vision: Essays in Modern Economic Perspectives*, eds. by James and Buchanan and Bettina Monissen, Frankfur/Mmain: Campus Verlag, 1998, 131-162.
- Chipman, J. S., and G. Tian, "Detrimental Externalities, Pollution Rights, and the "Coase Theorem'," *Economic Theory*, 49 (2012). 309-327.
- Coase, R., "The Problem of Social Cost," *Journal of Law and Economics*, 3 (1960), 1 44.
- Hurwicz, L., "What is the Coase Theorem," *Japan and the World Economy*, 7 (1995), 49-74.



- Laffont, J.-J., *Fundamentals of Public Economics*, Cambridge, MIT Press, 1988.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 9.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic*, Oxford University Press, 1995, Chapter 11.
- D. Meng and G. Tian, “Nonlinear Pricing with Network Externalities and Countervailing Incentives,” Texas A&M University, 2008, website: <http://econ.tamu.edu/tian/paper.htm>.
- Pigou, A., *A study of Public Finance*, New York: Macmillan, 1928.
- Salanie, B., *Microeconomics of Market Failures*, MIT Press, 2000, Chapter 6.
- Starrett, D. A., “Fundamental Nonconvexities in the Theory of Externalities,” *Journal of Economic Theory*, 4 (1972), 180-199.
- Tian, G. “Property Rights and the Nature of Chinese Collective Enterprises,” *Journal of Comparative Economics*, 28 (2000), 247-268.
- Tian, G. “A Theory of Ownership Arrangements and Smooth Transition To A Free Market Economy,” *Journal of Institutional and Theoretical Economics*, 157 (2001), 380-412. (The Chinese version published in *China Economic Quarterly*, 1 (2001), 45-70.)
- Tian, G., “A Solution to the Problem of Consumption Externalities,” *Journal of Mathematical Economics*, 39 (2003), 831-847.
- Tian, G., “A Unique Informationally Efficient Allocation Mechanism in Economies with Consumption Externalities,” *International Economic Review*, 45 (2004), 79-111.
- Tian, G. and L. Yang, “Theory of Negative Consumption Externalities with Applications to Economics of Happiness,” *Economic Theory*, 39 (2009), 399-424.
- Tian, G. and L. Yang, “Balanced Growth: A  $\bar{c}\bar{c}\bar{c}$  Potential Resolution to the Easterlin Paradox,” Texas A&M University, 2012, website: <http://econ.tamu.edu/tian/paper.htm>.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 24.

Varian, "A Solution to the Problem of Externalities when Agents Are Well Informed,"  
*American Economic Review*, 84 (1994), 1278 1293.

# Chapter 12

## Public Goods

### 12.1 Introduction

A *pure public good* is a good in which consuming one unit of the good by an individual in no way prevents others from consuming the same unit of the good. Thus, the good is nonexcludable and non-rival.

A good is *excludable* if people can be excluded from consuming it. A good is *non-rival* if one person's consumption does not reduce the amount available to other consumers.

*Examples of Public Goods:* street lights, policemen, fire protection, highway system, national defence, flood-control project, public television and radio broadcast, public parks, and a public project.

*Local Public Goods:* when there is a location restriction for the service of a public good.

Even if the competitive market is an efficient social institution for allocating private goods in an efficient manner, it turns out that a private market is not a very good mechanism for allocating public goods.

### 12.2 Notations and Basic Settings

In a general setting of public goods economy that includes consumers, producers, private goods, and public goods.

Let

$n$ : the number of consumers.

$L$ : the number of private goods.

$K$ : the number of public goods.

$Z_i \subseteq \mathfrak{R}_+^L \times \mathfrak{R}_+^K$ : the consumption space of consumer  $i$ .

$Z \subseteq \mathfrak{R}_+^{nL} \times \mathfrak{R}_+^K$ : consumption space.

$x_i \in \mathfrak{R}_+^L$ : a consumption of private goods by consumer  $i$ .

$y \in \mathfrak{R}_+^K$ : a consumption/production of public goods.

$w_i \in \mathfrak{R}_+^L$ : the initial endowment of private goods for consumer  $i$ . For simplicity, it is assumed that there is no public goods endowment, but they can be produced from private goods by a firm.

$v \in \mathfrak{R}_+^L$ : the private goods input. For simplicity, assume there is only one firm to produce the public goods.

$f: \mathfrak{R}_+^L \rightarrow \mathfrak{R}_+^K$ : production function with  $y = f(v)$ .

$\theta_i$ : the profit share of consumer  $i$  from the production.

$(x_i, y) \in Z_i$ .

$(x, y) = (x_i, \dots, x_n, y) \in Z$ : an allocation.

$\succsim_i$  (or  $u_i$  if exists) is a preference ordering.

$e_i = (Z_i, \succsim_i, w_i, \theta_i)$ : the characteristic of consumer  $i$ .

An allocation  $z \equiv (x, y)$  is feasible if

$$\sum_{i=1}^n x_i + v \leq \sum_{i=1}^n w_i \quad (12.1)$$

and

$$y = f(v) \quad (12.2)$$

$e = (e_1, \dots, e_n, f)$ : a public goods economy.

An allocation  $(x, y)$  is *Pareto efficient* for a public goods economy  $e$  if it is feasible and there is no other feasible allocation  $(x', y')$  such that  $(x'_i, y') \succsim_i (x_i, y)$  for all consumers  $i$  and  $(x'_k, y') \succ_k (x_k, y)$  for some  $k$ .

An allocation  $(x, y)$  is *weakly Pareto efficient* for the public goods economy  $e$  if it is feasible and there is no other feasible allocation  $(x', y')$  such that  $(x'_i, y') \succ_i (x_i, y)$  for all consumers  $i$ .

**Remark 12.2.1** Unlike private goods economies, even though under the assumptions of continuity and strict monotonicity, a weakly Pareto efficient allocation may not be Pareto efficient for the public goods economies. The following proposition is due to Tian (Economics Letters, 1988).

**Proposition 12.2.1** *For the public goods economies, a weakly Pareto efficient allocation may not be Pareto efficient even if preferences satisfy strict monotonicity and continuity.*

Proof: The proof is by way of an example. Consider an economy with  $(n, L, K) = (3, 1, 1)$ , constant returns in producing  $y$  from  $x$  (the input-output coefficient normalized to one), and the following endowments and utility functions:  $w_A = w_B = w_3 = 1$ ,  $u_1(x_1, y) = x_1 + y$ , and  $u_i(x_i, y) = x_i + 2y$  for  $i = 2, 3$ . Then  $z = (x, y)$  with  $x = (0.5, 0, 0)$  and  $y = 2.5$  is weakly Pareto efficient but not Pareto efficient because  $z' = (x', y') = (0, 0, 0, 3)$  Pareto-dominates  $z$  by consumers 2 and 3.

However, under an additional condition of strict convexity, they are equivalent. The proof is left to readers.

## 12.3 Discrete Public Goods

### 12.3.1 Efficient Provision of Public Goods

For simplicity, consider a public good economy with  $n$  consumers and two goods: one private good and one public good.

Let  $g_i$  be the contribution made by consumer  $i$ , so that

$$\begin{aligned} x_i + g_i &= w_i \\ \sum_{i=1}^n g_i &= v \end{aligned}$$

Assume  $u_i(x_i, y)$  is strictly monotonic increasing and continuous.

Let  $c$  be the cost of producing the public project so that the production technology is given by

$$y = \begin{cases} 1 & \text{if } \sum_{i=1}^n g_i \geq c \\ 0 & \text{otherwise} \end{cases}.$$

We first want to know under what conditions providing the public good will be Pareto dominate to not producing it, i.e., there exists  $(g_1, \dots, g_n)$  such that  $\sum_{i=1}^n g_i \geq c$  and

$$u_i(w_i - g_i, 1) > u_i(w_i, 0) \quad \forall i.$$

Let  $r_i$  be the maximum willingness-to-pay (reservation price) of consumer  $i$ , i.e.,  $r_i$  must satisfy

$$u_i(w_i - r_i, 1) = u_i(w_i, 0). \quad (12.3)$$

If producing the public project Pareto dominates not producing the public project, we have

$$u_i(w_i - g_i, 1) > u_i(w_i, 0) = u_i(w_i - r_i, 1) \quad \text{for all } i \quad (12.4)$$

By monotonicity of  $u_i$ , we have

$$w_i - g_i > w_i - r_i \quad \text{for } i \quad (12.5)$$

Then, we have

$$r_i > g_i \quad (12.6)$$

and thus

$$\sum_{i=1}^n r_i > \sum_{i=1}^n g_i \geq c \quad (12.7)$$

That is, the sum of the willingness-to-pay for the public good must exceed the cost of providing it. This condition is necessary. In fact, this condition is also sufficient. In summary, we have the following proposition.

**Proposition 12.3.1** *Providing a public good Pareto dominates not producing the public good if and only if  $\sum_{i=1}^n r_i > \sum_{i=1}^n g_i \geq c$ .*

### 12.3.2 Free-Rider Problem

How effective is a private market at providing public goods? The answer as shown below is that we cannot expect that purely independent decision will necessarily result in an

efficient amount of the public good being produced. To see this, suppose

$$\begin{aligned} r_i &= 100 \quad i = 1, 2 \\ c &= 150 \text{ (total cost)} \\ g_i &= \begin{cases} 150/2 = 75 & \text{if both agents make contributions} \\ 150 & \text{if only agent } i \text{ makes contribution} \end{cases} \end{aligned}$$

Each person decides independently whether or not to buy the public good. As a result, each one has an incentive to be a free-rider on the other as shown the following payoff matrix.

|          |             | Person 2   |             |
|----------|-------------|------------|-------------|
|          |             | Buy        | Doesn't Buy |
| Person 1 | Buy         | (25, 25)   | (-50, 100)  |
|          | Doesn't Buy | (100, -50) | (0, 0)      |

**Table 12.3.2: Private Provision of a Discrete Public Good**

Note that net payoffs are defined by  $r_i - g_i$ . Thus, it is given by  $100 - 150/2 = 25$  when both consumers are willing to produce the public project, and  $100 - 150 = -50$  when only one person wants to buy, but the other person does not.

This is the prisoner's dilemma. The dominant strategy equilibrium in this game is (doesn't buy, doesn't buy). Thus, no body wants to share the cost of producing the public project, but wants to free-ride on the other consumer. As a result, the public good is not provided at all even though it would be efficient to do so. Thus, voluntary contribution in general does not result in the efficient level of the public good.

### 12.3.3 Voting for a Discrete Public Good

The amount of a public good is often determined by a voting. Will this generally result in an efficient provision? The answer is no.

Voting does not result in efficient provision. Consider the following example.

### Example 12.3.1

$$c = 99$$

$$r_1 = 90, r_2 = 30, r_3 = 30$$

Clearly,  $r_1 + r_2 + r_3 > c$ .  $g_i = 99/3 = 33$ . So the efficient provision of the public good should be yes. However, under the majority rule, only consumer 1 votes “yes” since she receives a positive net benefit if the good is provided. The 2nd and 3rd persons vote “no” to produce public good, and therefore, the public good will not be provided so that we have inefficient provision of the public good. The problem with the majority rule is that it only measures the net benefit for the public good, whereas the efficient condition requires a comparison of willingness-to-pay.

## 12.4 Continuous Public Goods

### 12.4.1 Efficient Provision of Public Goods

Again, for simplicity, we assume there is only one public good and one private good that may be regarded as money, and  $y = f(v)$ .

The welfare maximization approach shows that Pareto efficient allocations can be characterized by

$$\begin{aligned} & \max_{(x,y)} \sum_{i=1}^n a_i u_i(x_i, y) \\ \text{s.t.} \quad & \sum_{i=1}^n x_i + v \leq \sum_{i=1}^n w_i \\ & y \leq f(v) \end{aligned}$$

Define the Lagrangian function:

$$L = \sum_{i=1}^n a_i u_i(x_i, y) + \lambda \left( \sum_{i=1}^n w_i - \sum_{i=1}^n x_i - v \right) + \mu(f(v) - y). \quad (12.8)$$

When  $u_i$  is strictly quasi-concave and differentiable and  $f(v)$  is concave and differentiable,



the set of Pareto optimal allocations are characterized by the first order condition:

$$\frac{\partial L}{\partial x_i} = 0 : \quad a_i \frac{\partial u_i}{\partial x_i} = \lambda \quad (12.9)$$

$$\frac{\partial L}{\partial v} = 0 : \quad \mu f'(v) = \lambda \quad (12.10)$$

$$\frac{\partial L}{\partial y} = 0 : \quad \sum_{i=1}^n a_i \frac{\partial u_i}{\partial y} = \mu. \quad (12.11)$$

So at an interior solution, by (12.9) and (12.10)

$$\frac{a_i}{\mu} = \frac{f'(v)}{\frac{\partial u_i}{\partial x_i}} \quad (12.12)$$

Substituting (12.12) into (12.11),

$$\sum_{i=1}^n \frac{\frac{\partial u_i}{\partial y}}{\frac{\partial u_i}{\partial x_i}} = \frac{1}{f'(v)}. \quad (12.13)$$

Thus, we obtain the well-known Lindahl-Samuelson condition.

In conclusion, the conditions for Pareto efficiency are given by

$$\begin{cases} \sum_{i=1}^n MRS_{yx_i}^i = MRTS_{yv} \\ \sum x_i + v \leq \sum_{i=1}^n w_i \\ y = f(v) \end{cases} \quad (12.14)$$

#### Example 12.4.1

$$u_i = a_i \ln y + \ln x_i$$

$$y = v$$

the Lindahl-Samuelson condition is

$$\sum_{i=1}^n \frac{\frac{\partial u_i}{\partial y}}{\frac{\partial u_i}{\partial x_i}} = 1 \quad (12.15)$$

and thus

$$\sum_{i=1}^n \frac{\frac{a_i}{y}}{\frac{1}{x_i}} = \sum_{i=1}^n \frac{a_i x_i}{y} = 1 \Rightarrow \sum a_i x_i = y \quad (12.16)$$

which implies the level of the public good is not uniquely determined.

Thus, in general, the marginal willingness-to-pay for a public good depends on the amount of private goods consumption, and therefor, the efficient level of  $y$  depends on  $x_i$ . However, in the case of quasi-linear utility functions,

$$u_i(x_i, y) = x_i + u_i(y) \quad (12.17)$$

the Lindahl-Samuelson condition becomes

$$\sum_{i=1}^n u'_i(y) = \frac{1}{f'(v)} \equiv c'(y) \quad (12.18)$$

and thus  $y$  is uniquely determined.

#### Example 12.4.2

$$\begin{aligned} u_i &= a_i \ln y + x_i \\ y &= v \end{aligned}$$

the Lindahl-Samuelson condition is

$$\sum_{i=1}^n \frac{\frac{\partial u_i}{\partial y}}{\frac{\partial u_i}{\partial x_i}} = 1 \quad (12.19)$$

and thus

$$\sum_{i=1}^n \frac{a_i}{y} = \sum_{i=1}^n \frac{a_i}{y} = 1 \Rightarrow \sum a_i = y \quad (12.20)$$

which implies the level of the public good is uniquely determined.

### 12.4.2 Lindahl Equilibrium

We have given the conditions for Pareto efficiency in the presence of public goods. The next problem is how to achieve a Pareto efficient allocation in a decentralized way. In private-goods-only economies, any competitive equilibrium is Pareto optimal. However, with public goods, a competitive mechanism does not help. For instance, if we tried the competitive solution with a public good and two consumers, the utility maximization would equalize the MRS and the relative price, e.g.,

$$MRS_{yx}^A = MRS_{yx}^B = \frac{p_y}{p_x}.$$

This is an immediate violation to the Samuelson-Lindahl optimal condition.

Lindahl suggested to use a tax method to provide a public good. Each person is signed a specific “personalized price” for the public good. The Lindahl solution is a way to mimic the competitive solution in the presence of public goods. Suppose we devise a mechanism to allocate the production cost of a public good between consumers, then we can achieve the Samuelson-Lindahl condition. To this end, we apply different prices of a public good

to different consumers. The idea of the Lindahl solution is that the consumption level of a public goods is the same to all consumers, but the price of the public good is personalized among consumers in the way that the price ratio of two goods for each person being equal the marginal rates of substitutions of two goods.

To see this, consider a public goods economy  $e$  with  $x_i \in R_+^L$  (private goods) and  $y \in R_+^K$  (public goods). For simplicity, we assume the CRS for  $y = f(v)$ . A feasible allocation

$$\sum_{i=1}^n x_i + v \leq \sum_{i=1}^n w_i \quad (12.21)$$

Let  $q_i \in R_+^K$  be the personalized price vector of consumer  $i$  for consuming the public goods.

Let  $q = \sum_{i=1}^n q_i$  : the market price vector of  $y$ .

Let  $p \in R_+^L$  be the price vector of private goods.

The profit is defined as  $\pi = qy - pv$  with  $y = f(v)$ .

**Definition 12.4.1 (Lindahl Equilibrium)** An allocation  $(x^*, y^*) \in R_+^{nL+K}$  is a Lindahl equilibrium allocation if it is feasible and there exists a price vector  $p^* \in R_+^L$  and personalized price vectors  $q_i^* \in R_+^K$ , one for each individual  $i$ , such that

- (i)  $p^*x_i^* + q_i^*y^* \leq p^*w_i$ ;
- (ii)  $(x_i, y) \succ_i (x_i^*, y^*)$  implies  $p^*x_i + q_i^*y > p^*w_i$ ;
- (iii)  $q^*y^* - p^*v^* = 0$ ,

where  $v^* = \sum_{i=1}^n w_i - \sum_{i=1}^n x_i^*$  and  $\sum_{i=1}^n q_i^* = q^*$ .

We call  $(x^*, y^*, p^*, q_1^*, \dots, q_n^*)$  a Lindahl equilibrium.

**Remark 12.4.1** Because of CRS, the maximum profit is zero at the Lindahl equilibrium.

We may regard a Walrasian equilibrium as a special case of a Lindahl equilibrium when there are no public goods. In fact, the concept of Lindahl equilibrium in economies with public goods is, in many ways, a natural generalization of the Walrasian equilibrium notion in private goods economies, with attention to the well-known duality that reverses the role of prices and quantities between private and public goods, and between Walrasian and Lindahl allocations. In the Walrasian case, prices must be equalized while quantities

are individualized; in the Lindahl case the quantities of the public good must be the same for everyone, while prices charged for public goods are individualized. In addition, the concepts of Walrasian and Lindahl equilibria are both relevant to private-ownership economies. Furthermore, they are characterized by purely price-taking behavior on the part of agents. It is essentially this property that one can consider the Lindahl solution as an informationally decentralized process.

Similarly, we have the following First Fundamental Theorem of Welfare Economics for public goods economies.

**Theorem 12.4.1** : *Every Lindahl allocation  $(x^*, y^*)$  with the price system  $(p^*, q_1^*, \dots, q_n^*)$  is weakly Pareto efficient, and further under local non-satiation, it is Pareto efficient.*

*Proof:* We only prove the second part. The first part is the simple. Suppose not. There exists another feasible allocation  $(x_i, y)$  such that  $(x_i, y) \succsim_i (x_i^*, y^*)$  for all  $i$  and  $(x_j, y) \succ_j (x_j^*, y^*)$  for some  $j$ . Then, by local-non-satiation of  $\succsim_i$ , we have

$$\begin{aligned} p^* x_i + q_i^* y &\geq p^* w_i & \text{for all } i = 1, 2, \dots, n \\ p^* x_j + q_j^* y &> p^* w_j & \text{for some } j. \end{aligned}$$

Thus

$$\sum_{i=1}^n p^* x_i + \sum_{i=1}^n q_i y > \sum_{i=1}^n p^* w_i \quad (12.22)$$

So

$$p^* \sum_{i=1}^n x_i + q^* y > \sum_{i=1}^n p^* w_i$$

or

$$p^* \sum_{i=1}^n x_i + p^* v > \sum_{i=1}^n p^* w_i$$

by noting that  $q^* y - p^* v \leq q^* y^* - p^* v^* = 0$ . Hence,

$$p^* \left[ \sum_{i=1}^n (x_i - w_i) + v \right] > 0$$

which contradicts the fact that  $(x, y)$  is feasible.

For a public economy with one private good and one public good  $y = \frac{1}{q}v$ , the definition of Lindahl equilibrium becomes much simpler.

An allocation  $(x^*, y^*)$  is a Lindahl Allocation if  $(x^*, y^*)$  is feasible (i.e.,  $\sum_{i=1}^n x_i^* + qy^* \leq \sum w_i$ ) and there exists  $q_i^*$ ,  $i = 1, \dots, n$  such that

- (i)  $x_i^* + q_i y^* \leq w_i$
- (ii)  $(x_i, y) \succ_i (x_i^*, y^*)$  implies  $x_i + q_i y > w_i$
- (iii)  $\sum_{i=1}^n q_i = q$

In fact, the feasibility condition is automatically satisfied when the budget constraints (i) is satisfied.

If  $(x^*, y^*)$  is an interior Lindahl equilibrium allocation, from the utility maximization, we can have the first order condition:

$$\frac{\frac{\partial u_i}{\partial y}}{\frac{\partial u_i}{\partial x_i}} = \frac{q_i}{1} \quad (12.23)$$

which means the Lindahl-Samuelson condition holds:

$$\sum_{i=1}^n MRS_{yx_i} = q,$$

which is the necessary condition for Pareto efficiency.

### Example 12.4.3

$$\begin{aligned} u_i(x_i, y) &= x_i^{\alpha_i} y^{(1-\alpha_i)} \quad \text{for } 0 < \alpha_i < 1 \\ y &= \frac{1}{q} v \end{aligned}$$

The budget constraint is:

$$x_i + q_i y = w_i.$$

The demand functions for  $x_i$  and  $y_i$  of each  $i$  are given by

$$x_i = \alpha_i w_i \quad (12.24)$$

$$y_i = \frac{(1 - \alpha_i) w_i}{q_i} \quad (12.25)$$

Since  $y_1 = y_2 = \dots y_n = y^*$  at the equilibrium, we have by (12.25)

$$q_i y^* = (1 - \alpha_i) w_i. \quad (12.26)$$

Making summation, we have

$$q y^* = \sum_{i=1}^n (1 - \alpha_i) w_i.$$

Then, we have

$$y^* = \frac{\sum_{i=1}^n (1 - \alpha_i) w_i}{q}$$

and thus, by (12.26), we have

$$q_i = \frac{(1 - \alpha_i) w_i}{y^*} = \frac{q(1 - \alpha_i) w_i}{\sum_{i=1}^n (1 - \alpha_i) w_i}. \quad (12.27)$$

If we want to find a Lindahl equilibrium, we must know the preferences or MRS of each consumer.

But because of the free-rider problem, it is very difficult for consumers to report their preferences truthfully.

### 12.4.3 Free-Rider Problem

When the MRS is known, a Pareto efficient allocation  $(x, y)$  can be determined from the Lindahl-Samuelson condition or the Lindahl solution. After that, the contribution of each consumer is given by  $g_i = w_i - x_i$ . However, the society is hard to know the information about MRS. Of course, a naive method is that we could ask each individual to reveal his preferences, and thus determine the willingness-to-pay. However, since each consumer is self-interested, each person wants to be a free-rider and thus is not willing to tell the true MRS. If consumers realize that shares of the contribution for producing public goods (or the personalized prices) depend on their answers, they have “incentives to cheat.” That is, when the consumers are asked to report their utility functions or MRSs, they will have incentives to report a smaller *MRS* so that they can pay less, and consume the public good (free riders). This causes the major difficulty in the public economies.

To see this, notice that the social goal is to reach Pareto efficient allocations for the public goods economy, but from the personal interest, each person solves the following problem:

$$\max u_i(x_i, y) \quad (12.28)$$

subject to

$$\begin{aligned} g_i &\in [0, w_i] \\ x_i + g_i &= w_i \\ y &= f(g_i + \sum_{j \neq i}^n g_j). \end{aligned}$$

That is, each consumer  $i$  takes others' strategies  $g_{-i}$  as given, and maximizes his payoffs. From this problem, we can form a non-cooperative game:

$$\Gamma = (G_i, \phi_i)_{i=1}^n$$

where  $G_i = [0, w_i]$  is the strategy space of consumer  $i$  and  $\phi_i : G_1 \times G_2 \times \dots \times G_n \rightarrow R$  is the payoff function of  $i$  which is defined by

$$\phi_i(g_i, g_{-i}) = u_i[(w_i - g_i), f(g_i + \sum_{j \neq i}^n g_j)] \quad (12.29)$$

**Definition 12.4.2** For the game,  $\Gamma = (G_i, \phi_i)_{i=1}^n$ , the strategy  $g^* = (g_1^*, \dots, g_n^*)$  is a *Nash Equilibrium* if

$$\phi_i(g_i^*, g_{-i}^*) \geq \phi_i(g_i, g_{-i}^*) \quad \text{for all } g_i \in G_i \text{ and all } i = 1, 2, \dots, n,$$

$g^*$  is a *dominant strategy equilibrium* if

$$\phi_i(g_i^*, g_{-i}) \geq \phi_i(g_i, g_{-i}) \quad \text{for all } g \in G \text{ and all } i = 1, 2, \dots$$

**Remark 12.4.2** Note that the difference between Nash equilibrium (NE) and dominant strategy is that at NE, given best strategy of others, each consumer chooses his best strategy while dominant strategy means that the strategy chosen by each consumer is best regardless of others' strategies. Thus, a dominant strategy equilibrium is clearly a Nash equilibrium, but the converse may not be true. Only for a very special payoff functions, there is a dominant strategy while a Nash equilibrium exists for a continuous and quasi-concave payoff functions that are defined on a compact set.

For Nash equilibrium, if  $u_i$  and  $f$  are differentiable, then the first order condition is:

$$\frac{\partial \phi_i(g^*)}{\partial g_i} \leq 0 \quad \text{with equality if } g_i > 0 \quad \text{for all } i = 1, \dots, n. \quad (12.30)$$

Thus, we have

$$\frac{\partial \phi_i}{\partial g_i} = \frac{\partial u_i}{\partial x_i}(-1) + \frac{\partial u_i}{\partial y} f'(g_i^* + \sum_{j \neq i}^n g_j) \leq 0 \quad \text{with equality if } g_i > 0.$$

So, at an interior solution  $g^*$ , we have

$$\frac{\frac{\partial u_i}{\partial y}}{\frac{\partial u_i}{\partial x_i}} = \frac{1}{f'(g_i^* + \sum_{j \neq i}^n g_j)},$$

and thus

$$MRS_{yx_i}^i = MRTS_{yv},$$

which does not satisfy the Lindahl-Samuelson condition. Thus, the Nash equilibrium in general does not result in Pareto efficient allocations. The above equation implies that the low level of public good is produced rather than the Pareto efficient level of the public good when utility functions are quasi-concave. Therefore, Nash equilibrium allocations are in general not consistent with Pareto efficient allocations. How can one solve this free-ride problem? We will answer this question in the mechanism design theory.

## Reference

- Foley, D., “Lindahl’s Solution and the Core of an Economy with Public Goods, *Econometrica* 38 (1970), 66–72.
- Laffont, J.-J., *Fundamentals of Public Economics*, Cambridge, MIT Press, 1988, Chapter 2.
- Lindahl, E., “Die Gerechtigkeit der Besteuerung. Lund: Gleerup.[English translation: Just Taxation – a Positive Solution: In *Classics in the Theory of Public Finance*, edited by R. A. Musgrave and A. T. Peacock. London: Macmillan, 1958].
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 9.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic*, Oxford University Press, 1995, Chapter 11.
- Milleron, J. C. “Theory of Value with Public Goods: a Survey Article,” *Journal of Economics Theory* 5 (1972), 419–477.
- Muench, T., “The Core and the Lindahl Equilibrium of an Economy with a Public Good, *Journal of Economics Theory* 4 (1972), 241–255.
- Pigou, A., *A study of Public Finance*, New York: Macmillan, 1928.
- Salanie, B., *Microeconomics of Market Failures*, MIT Press, 2000, Chapter 5.



- Roberts, D. J., “The Lindahl Solution for Economies with Public Goods, *Journal of Public Economics* 3 (1974), 23 42.
- Tian, G., “On the Constrained Walrasian and Lindahl Correspondences,” *Economics Letters* 26 (1988), 299 303.
- Tian, G. and Q. Li, “Ratio-Lindahl and Ratio Equilibria with Many Goods, *Games and Economic Behavior* 7 (1994), 441 460.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 23.
- Varian, “A Solution to the Problem of Externalities when Agents Are Well Informed,” *American Economic Review*, 84 (1994), 1278 1293.

## Part V

# Information, Incentives, Mechanism Design, and Contract Theory

Information economics, incentive theory, mechanism design theory, principal-agent theory, contract theory, and auction theory have been very important and active research areas and had wide applications in various fields in economics, finance, management, and corporate law, political sciences in last five decades. Because of this, about twenty economists of founding contributors of mechanism design and the associated files of game theory so far have been rewarded with the Nobel prize in economics, including Hayek, Ken Arrow, Gerard Debreu, Ronald Coase, Herbert Simon, John Nash, Reinhard Selten, William Vickrey, James Mirrlees, George Akerlof, Joseph Stiglitz, Michael Spence, Robert Auman, Leo Hurwicz, Eric Maskin, Roger Myerson, Peter Diamond, Oliver Williamson, Al Roth, and Lloyd S. Shapley.

The notion of incentives is a basic and key concept in modern economics. To many economists, economics is to a large extent a matter of incentives: incentives to work hard, to produce good quality products, to study, to invest, to save, etc.

Until about 40 year ago, economics was mostly concerned with understanding the theory of value in large economies. A central question asked in general equilibrium theory was whether a certain mechanism (especially the competitive mechanism) generated Pareto-efficient allocations, and if so – for what categories of economic environments. In a perfectly competitive market, the pressure of competitive markets solves the problem of incentives for consumers and producers. The major project of understanding how prices are formed in competitive markets can proceed without worrying about incentives.

The question was then reversed in the economics literature: instead of regarding mechanisms as given and seeking the class of environments for which they work, one seeks mechanisms which will implement some desirable outcomes (especially those which result in Pareto-efficient and individually rational allocations) for a given class of environments without destroying participants' incentives, and have a low cost of operation and other desirable properties. In a sense, the theorists went back to basics.

The reverse question was stimulated by two major lines in the history of economics. Within the capitalist/private-ownership economics literature, a stimulus arose from studies focusing upon the failure of the competitive market to function as a mechanism for implementing efficient allocations in many nonclassical economic environments such as the presence of externalities, public goods, incomplete information, imperfect compe-

tition, increasing return to scale, etc. At the beginning of the seventies, works by Akerlof (1970), Hurwicz (1972), Spence (1974), and Rothschild and Stiglitz (1976) showed in various ways that asymmetric information was posing a much greater challenge and could not be satisfactorily imbedded in a proper generalization of the Arrow-Debreu theory.

A second stimulus arose from the socialist/state-ownership economics literature, as evidenced in the “socialist controversy” — the debate between Mises-Hayek and Lange-Lerner in twenties and thirties of the last century. The controversy was provoked by von Mises’s skepticism as to even a theoretical feasibility of rational allocation under socialism.

The incentives structure and information structure are thus two basic features of any economic system. The study of these two features is attributed to these two major lines, culminating in the theory of mechanism design. The theory of economic mechanism design which was originated by Hurwicz is very general. All economic mechanisms and systems (including those known and unknown, private-ownership, state-ownership, and mixed-ownership systems) can be studied with this theory.

At the micro level, the development of the theory of incentives has also been a major advance in economics in the last forty years. Before, by treating the firm as a black box the theory remains silent on how the owners of firms succeed in aligning the objectives of its various members, such as workers, supervisors, and managers, with profit maximization.

When economists began to look more carefully at the firm, either in agricultural or managerial economics, incentives became the central focus of their analysis. Indeed, delegation of a task to an agent who has different objectives than the principal who delegates this task is problematic when information about the agent is imperfect. This problem is the essence of incentive questions. Thus, **conflicting objectives and decentralized information are the two basic ingredients of incentive theory.**

We will discover that, in general, these informational problems prevent society from achieving the first-best allocation of resources that could be possible in a world where all information would be common knowledge. The additional costs that must be incurred because of the strategic behavior of privately informed economic agents can be viewed as one category of the transaction costs. Although they do not exhaust all possible transaction costs, economists have been rather successful during the last thirty years in modeling and analyzing these types of costs and providing a good understanding of the

limits set by these on the allocation of resources. This line of research also provides a whole set of insights on how to begin to take into account agents' responses to the incentives provided by institutions.

The three words — contracts, mechanisms and institutions are to a large extent synonymous. They all mean “rules of the game,” which describe what actions the parties can undertake, and what outcomes these actions would be obtained. In most cases the rules of the game are designed by someone: in chess, basketball, etc. The rules are designed to achieve better outcomes. But there is one difference. While mechanism design theory may be able answer “big” questions, such as “socialism vs. capitalism,” contract theory is developed and useful for more manageable smaller questions, concerning specific contracting practices and mechanisms.

Thus, mechanism design is normative economics, in contrast to game theory, which is positive economics. Game theory is important because it predicts how a given game will be played by agents. Mechanism design goes one step further: given the physical environment and the constraints faced by the designer, what goal can be realized or implemented? What mechanisms are optimal among those that are feasible? In designing mechanisms one must take into account incentive constraints (e.g., consumers may not report truthfully how many pairs of shoes they need or how productive they are).

This part considers the design of contracts in which one or many parties have private characteristics or hidden actions. The party who designs the contract will be called the Principal, while the other parties will be called agents. For the most part we will focus on the situation where the Principal has no private information and the agents do. This framework is called *screening*, because the principal will in general try to screen different types of agents by inducing them to choose different bundles. The opposite situation, in which the Principal has private information and the agents do not, is called *signaling*, since the Principal could signal his type with the design of his contract.

We will briefly present the contract theory in four chapters. Chapters 13 and 14 consider the principal-agent model where the principal delegates an action to a single agent with private information. This private information can be of two types: either the agent can take an action unobserved by the principal, the case of moral hazard or hidden action; or the agent has some private knowledge about his cost or valuation that is ignored by

the principal, the case of adverse selection or hidden knowledge. The theory of optimal mechanism design considers when this private information is a problem for the principal, and what is the optimal way for the principal to cope with it. **The design of the principal's optimal contract can be regarded as a simple optimization problem. This simple focus will turn out to be enough to highlight the various trade-offs between allocative efficiency and distribution of information rents arising under incomplete information.** The mere existence of informational constraints may generally prevent the principal from achieving allocative efficiency. We will characterize the allocative distortions that the principal finds desirable to implement in order to mitigate the impact of informational constraints.

Chapter 15 will consider situations with one principal and many agents. Asymmetric information may not only affect the relationship between the principal and each of his agents, but it may also plague the relationships between agents. Moreover, maintaining the hypothesis that agents adopt an individualistic behavior, those organizational contexts require a solution concept of equilibrium, which describes the strategic interaction between agents under complete or incomplete information.

Chapter 16 will briefly study dynamic contract theory. We will discuss long-term incentive contracting in a dynamic principal-agent setting with one-agent and adverse selection. We will first consider the case where the principal (designer) can commit to a contract forever, and then consider what happens when she cannot commit against modifying the contract as new information arrives.

It is greatly acknowledged that the materials are largely drawn from the references in the end of each chapter, especially from Laffont and Martimort (1992) for chapters 13 and 14 and from Segal (2007).

# Chapter 13

## Optimal Mechanism Design: Contracts with One-Agent and Hidden Information

### 13.1 Introduction

The optimal contract theory with two parties in large extent is also called principal-agent theory. One party is a principal and the other is an agent.

Incentive problems arise when a principal wants to delegate a task to an agent with private information. The exact opportunity cost of this task, the precise technology used, and how good the matching is between the agent's intrinsic ability and this technology are all examples of pieces of information that may become private knowledge of the agent. In such cases, we will say that there is adverse selection.

#### Examples

1. The landlord delegates the cultivation of his land to a tenant, who will be the only one to observe the exact local weather conditions.
2. A client delegates his defense to an attorney who will be the only one to know the difficulty of the case.
3. An investor delegates the management of his portfolio to a broker, who will privately know the prospects of the possible investments.
4. A stockholder delegates the firm's day-to-day decisions to a manager, who will be

the only one to know the business conditions.

5. An insurance company provides insurance to agents who privately know how good a driver they are.

6. The Department of Defense procures a good from the military industry without knowing its exact cost structure.

7. A regulatory agency contracts for service with a public utility company without having complete information about its technology.

The common aspect of all those contracting settings is that the information gap between the principal and the agent has some fundamental implications for the design of the contract they sign. In order to reach an efficient use of economic resources, some information rent must be given up to the privately informed agent. At the optimal second-best contract, the principal trades off his desire to reach allocative efficiency against the costly information rent given up to the agent to induce information revelation. Implicit here is the idea that there exists a legal framework for this contractual relationship. The contract can be enforced by a benevolent court of law, the agent is bounded by the terms of the contract.

The main objective of this chapter is to characterize the optimal rent extraction-efficiency trade-off faced by the principal when designing his contractual offer to the agent under the set of incentive feasible constraints: incentive and participation constraints. In general, incentive constraints are binding at the optimum, showing that adverse selection clearly impedes the efficiency of trade. **The main lessons of this optimization is that the optimal second-best contract calls for a distortion in the volume of trade away from the first-best and for giving up some strictly positive information rents to the most efficient agents.**



## 13.2 Basic Settings of Principal-Agent Model with Adverse Selection

### 13.2.1 Economic Environment (Technology, Preferences, and Information)

Consider a consumer or a firm (the principal) who wants to delegate to an agent the production of  $q$  units of a good. The value for the principal of these  $q$  units is  $S(q)$  where  $S' > 0$ ,  $S'' < 0$  and  $S(0) = 0$ .

The production cost of the agent is unobservable to the principal, but it is common knowledge that the fixed cost is  $F$  and the marginal cost belongs to the set  $\Phi = \{\underline{\theta}, \bar{\theta}\}$ . The agent can be either efficient ( $\underline{\theta}$ ) or inefficient ( $\bar{\theta}$ ) with respective probabilities  $\nu$  and  $1 - \nu$ . That is, he has the cost function

$$C(q, \underline{\theta}) = \underline{\theta}q + F \quad \text{with probability } \nu \quad (13.1)$$

or

$$C(q, \bar{\theta}) = \bar{\theta}q + F \quad \text{with probability } 1 - \nu \quad (13.2)$$

Denote by  $\Delta\theta = \bar{\theta} - \underline{\theta} > 0$  the spread of uncertainty on the agent's marginal cost. This information structure is exogenously given to the players.

### 13.2.2 Contracting Variables: Outcomes

The contracting variables are the quantity produced  $q$  and the transfer  $t$  received by the agent. Let  $\mathcal{A}$  be the set of feasible allocations that is given by

$$\mathcal{A} = \{(q, t) : q \in \mathbb{R}_+, t \in \mathbb{R}\} \quad (13.3)$$

These variables are both observable and verifiable by a third party such as a benevolent court of law.

### 13.2.3 Timing

Unless explicitly stated, we will maintain the timing defined in the figure below, where  $A$  denotes the agent and  $P$  the principal.

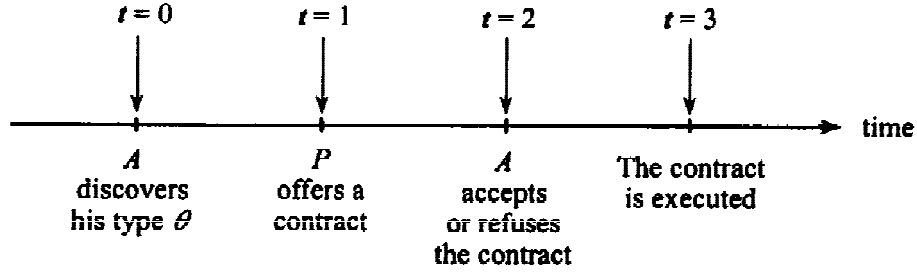


Figure 13.1: Timing of contracting under hidden information.

Note that contracts are offered at the interim stage; there is already asymmetric information between the contracting parties when the principal makes his offer.

### 13.3 The Complete Information Optimal Contract (Benchmark Case)

#### 13.3.1 First-Best Production Levels

To get a reference system for comparison, let us first suppose that there is no asymmetry of information between the principal and the agent. The efficient production levels are obtained by equating the principal's marginal value and the agent's marginal cost. Hence, we have the following first-order conditions

$$S'(\underline{q}^*) = \underline{\theta} \quad (13.4)$$

and

$$S'(\bar{q}^*) = \bar{\theta}. \quad (13.5)$$

The complete information efficient production levels  $\underline{q}^*$  and  $\bar{q}^*$  should be both carried out if their social values, respectively  $\underline{W}^* = S(\underline{q}^*) - \underline{\theta}\underline{q}^* - F$ , and  $\bar{W}^* = S(\bar{q}^*) - \bar{\theta}\bar{q}^* - F$ , are non-negative.

Since

$$S(\underline{q}^*) - \underline{\theta}\underline{q}^* \geq S(\bar{q}^*) - \underline{\theta}\bar{q}^* \geq S(\bar{q}^*) - \bar{\theta}\bar{q}^*$$

by definition of  $\underline{\theta}$  and  $\bar{\theta} > \underline{\theta}$ , the social value of production when the agent is efficient,  $\underline{W}^*$ , is greater than when he is inefficient, namely  $\bar{W}^*$ .

For trade to be always carried out, it is thus enough that production be socially valuable for the least efficient type, i.e., the following condition must be satisfied

$$\bar{W}^* = S(\bar{q}^*) - \bar{\theta}\bar{q}^* - F \geq 0. \quad (13.6)$$

As the fixed cost  $F$  plays no role other than justifying the existence of a single agent, it is set to zero from now on in order to simplify notations.

Note that, since the principal's marginal value of output is decreasing, the optimal production of an efficient agent is greater than that of an inefficient agent, i.e.,  $\underline{q}^* > \bar{q}^*$ .

### 13.3.2 Implementation of the First-Best

For a successful delegation of the task, the principal must offer the agent a utility level that is at least as high as the utility level that the agent obtains outside the relationship. We refer to these constraints as the agent's participation constraints. If we normalize to zero the agent's outside opportunity utility level (sometimes called his quo utility level), these participation constraints are written as

$$\underline{t} - \underline{\theta}\underline{q} \geq 0, \quad (13.7)$$

$$\bar{t} - \bar{\theta}\bar{q} \geq 0. \quad (13.8)$$

To implement the first-best production levels, the principal can make the following take-it-or-leave-it offers to the agent: If  $\theta = \bar{\theta}$  (resp.  $\underline{\theta}$ ), the principal offers the transfer  $\bar{t}^*$  (resp.  $\underline{t}^*$ ) for the production level  $\bar{q}^*$  (resp.  $\underline{q}^*$ ) with  $\bar{t}^* = \bar{\theta}\bar{q}^*$  (resp.  $\underline{t}^* = \underline{\theta}\underline{q}^*$ ). Thus, whatever his type, the agent accepts the offer and makes zero profit. The complete information optimal contracts are thus  $(\underline{t}^*, \underline{q}^*)$  if  $\theta = \underline{\theta}$  and  $(\bar{t}^*, \bar{q}^*)$  if  $\theta = \bar{\theta}$ . Importantly, under complete information delegation is costless for the principal, who achieves the same utility level that he would get if he was carrying out the task himself (with the same cost function as the agent).

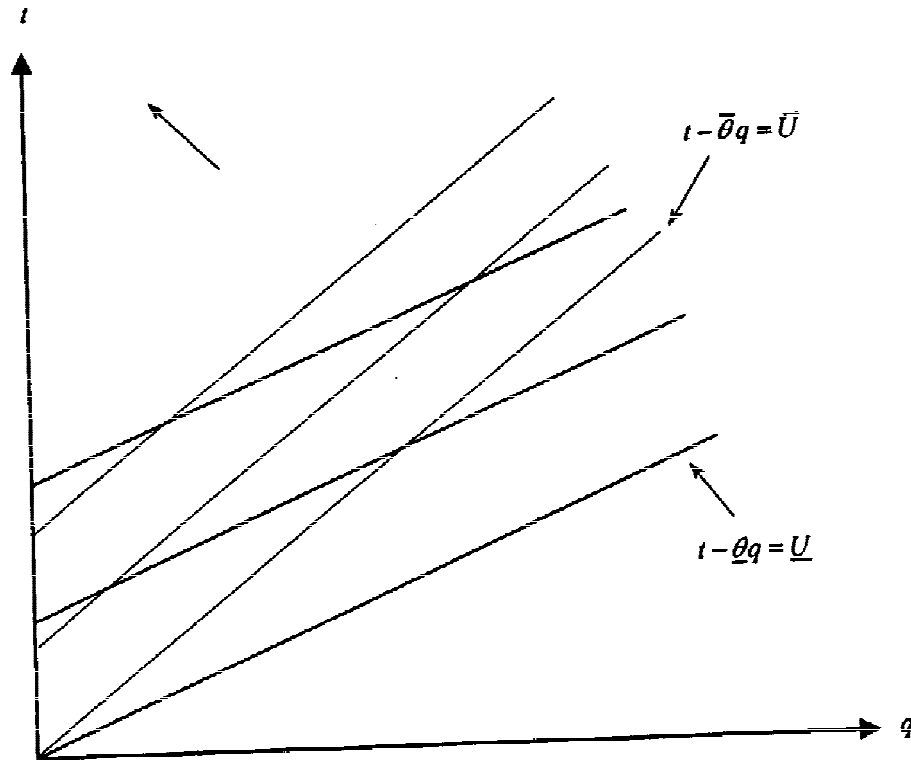


Figure 13.2: Indifference curves of both types.

### 13.3.3 A Graphical Representation of the Complete Information Optimal Contract

Since  $\bar{\theta} > \underline{\theta}$ , the iso-utility curves for different types cross only once as shown in the above figure. This important property is called the single-crossing or Spence-Mirrlees property.

The complete information optimal contract is finally represented Figure 13.3 by the pair of points  $(A^*, B^*)$ . Note that since the iso-utility curves of the principal correspond to increasing levels of utility when one moves in the southeast direction, the principal reaches a higher profit when dealing with the efficient type. We denote by  $\bar{V}^*$  (resp.  $\underline{V}^*$ ) the principal's level of utility when he faces the  $\bar{\theta}$ - (resp.  $\underline{\theta}$ -) type. Because the principal's has all the bargaining power in designing the contract, we have  $\bar{V}^* = \bar{W}^*$  (resp.  $\underline{V}^* = \underline{W}^*$ ) under complete information.

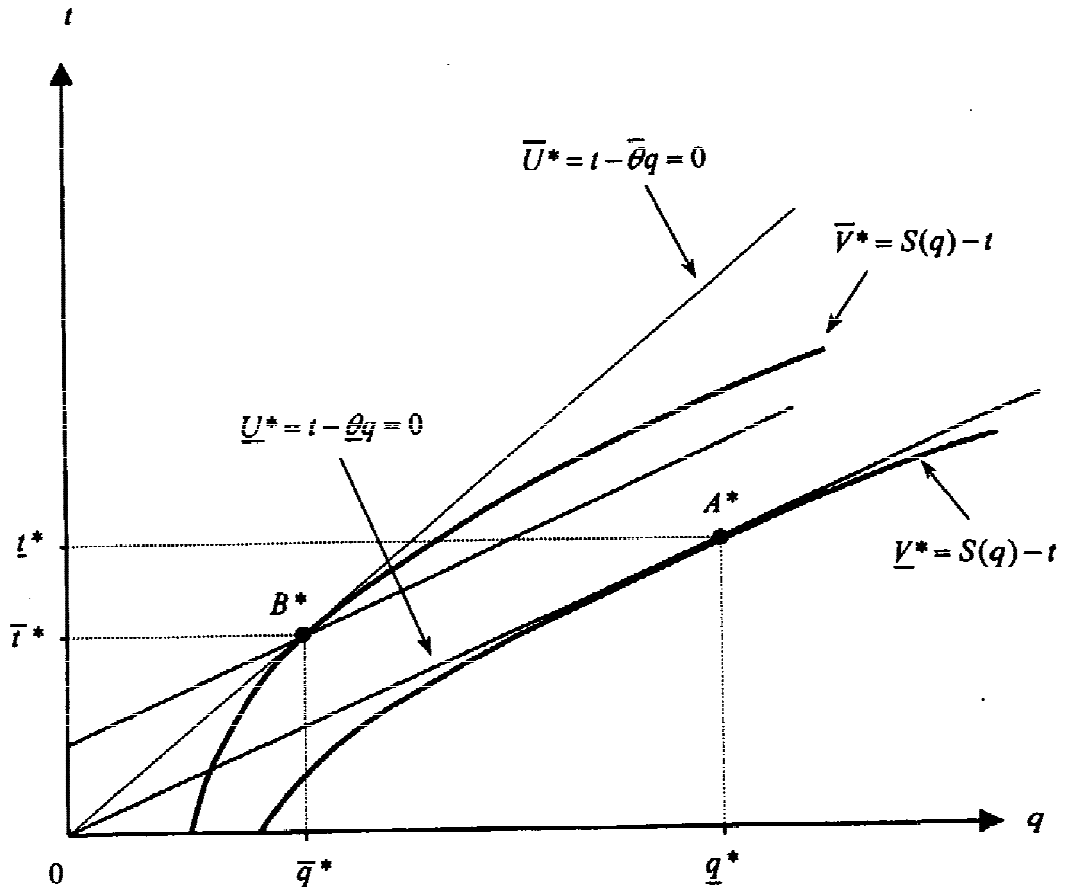


Figure 13.3: First best contracts.

## 13.4 Incentive Feasible Contracts

### 13.4.1 Incentive Compatibility and Participation

Suppose now that the marginal cost  $\theta$  is the agent's private information and let us consider the case where the principal offers the menu of contracts  $\{(t^*, q^*); (\bar{t}^*, \bar{q}^*)\}$  hoping that an agent with type  $\underline{\theta}$  will select  $(t^*, q^*)$  and an agent with  $\bar{\theta}$  will select instead  $(\bar{t}^*, \bar{q}^*)$ .

From Figure 13.3 above, we see that  $B^*$  is preferred to  $A^*$  by both types of agents. Offering the menu  $(A^*, B^*)$  fails to have the agents self-selecting properly within this menu. The efficient type have incentives to mimic the inefficient one and selects also contract  $B^*$ . The complete information optimal contracts can no longer be implemented under asymmetric information. We will thus say that the menu of contracts  $\{(t^*, q^*); (\bar{t}^*, \bar{q}^*)\}$  is not incentive compatible.

**Definition 13.4.1** A menu of contracts  $\{(\underline{t}, \underline{q}); (\bar{t}, \bar{q})\}$  is incentive compatible when  $(\underline{t}, \underline{q})$  is weakly preferred to  $(\bar{t}, \bar{q})$  by agent  $\underline{\theta}$  and  $(\bar{t}, \bar{q})$  is weakly preferred to  $(\underline{t}, \underline{q})$  by agent  $\bar{\theta}$ .

Mathematically, these requirements amount to the fact that the allocations must satisfy the following *incentive compatibility constraints*:

$$\underline{t} - \underline{\theta}\underline{q} \geq \bar{t} - \underline{\theta}\bar{q} \quad (13.9)$$

and

$$\bar{t} - \bar{\theta}\bar{q} \geq \underline{t} - \bar{\theta}\underline{q} \quad (13.10)$$

Furthermore, for a menu to be accepted, it must satisfy the following two *participation constraints*:

$$\underline{t} - \underline{\theta}\underline{q} \geq 0, \quad (13.11)$$

$$\bar{t} - \bar{\theta}\bar{q} \geq 0. \quad (13.12)$$

**Definition 13.4.2** A menu of contracts is incentive feasible if it satisfies both incentive and participation constraints (13.9) through (13.12).

The inequalities (13.9) through (13.12) fully characterize the set of incentive feasible menus of contracts. The restrictions embodied in this set express additional constraints imposed on the allocation of resources by asymmetric information between the principal and the agent.

### 13.4.2 Special Cases

**Bunching or Pooling Contracts:** A first special case of incentive feasible menu of contracts is obtained when the contracts targeted for each type coincide, i.e., when  $\underline{t} = \bar{t} = t^p, \underline{q} = \bar{q} = q^p$  and both types of agent accept this contract.

**Shutdown of the Least Efficient Type:** Another particular case occurs when one of the contracts is the null contract  $(0,0)$  and the nonzero contract  $(t^s, q^s)$  is only accepted by the efficient type. Then, (13.9) and (13.11) both reduce to

$$t^s - \underline{\theta}q^s \geq 0. \quad (13.13)$$

The incentive constraint of the bad type reduces to

$$0 \geq t^s - \bar{\theta}q^s. \quad (13.14)$$

As with the pooling contract, the benefit of the  $(0,0)$  option is that it somewhat reduces the number of constraints since the incentive and participation constraints take the same form. The cost of such a contract may be an excessive screening of types. Here, the screening of types takes the rather extreme form of the least efficient type.

### 13.4.3 Monotonicity Constraints

Incentive compatibility constraints reduce the set of feasible allocations. Moreover, these quantities must generally satisfy a monotonicity constraint which does not exist under complete information. Adding (13.9) and (13.10), we immediately have

$$\underline{q} \geq \bar{q}. \quad (13.15)$$

We will call condition (13.15) an implementability condition that is necessary and sufficient for implementability.

Indeed, suppose that (13.15) holds; it is clear that there exists transfers  $\bar{t}$  and  $\underline{t}$  such that the incentive constraints (13.9) and (13.10) both hold. It is enough to take those transfers such that

$$\theta(\underline{q} - \bar{q}) \leq \underline{t} - \bar{t} \leq \bar{\theta}(\underline{q} - \bar{q}). \quad (13.16)$$

**Remark 13.4.1** In this two-type model, the conditions for implementability take a simple form. With more than two types (or with a continuum), the characterization of these conditions might get harder. The conditions for implementability are also more difficult to characterize when the agent performs several tasks on behalf of the principal.

## 13.5 Information Rents

To understand the structure of the optimal contract it is useful to introduce the concept of information rent.

We know from previous discussion, under complete information, the principal is able to maintain all types of agents at their zero status quo utility level. Their respective utility levels  $\underline{U}^*$  and  $\bar{U}^*$  at the first-best satisfy

$$\underline{U}^* = \underline{t}^* - \underline{\theta}\underline{q}^* = 0 \quad (13.17)$$

and

$$\bar{U}^* = \bar{t}^* - \bar{\theta}\bar{q}^* = 0. \quad (13.18)$$

Generally this will not be possible anymore under incomplete information, at least when the principal wants both types of agents to be active.

We use the notations  $\underline{U} = \underline{t} - \underline{\theta}\underline{q}$  and  $\bar{U} = \bar{t} - \bar{\theta}\bar{q}$  to denote the respective information rent of each type.

Take any menu  $\{(\bar{t}, \bar{q}); (\underline{t}, \underline{q})\}$  of incentive feasible contracts. How much would a  $\underline{\theta}$ -agent get by mimicking a  $\bar{\theta}$ -agent? The high-efficient agent would get

$$\bar{t} - \underline{\theta}\bar{q} = \bar{t} - \bar{\theta}\bar{q} + \Delta\theta\bar{q} = \bar{U} + \Delta\theta\bar{q}. \quad (13.19)$$

Thus, as long as the principal insists on a positive output for the inefficient type,  $\bar{q} > 0$ , the principal must give up a positive rent to a  $\underline{\theta}$ -agent. This information rent is generated by the informational advantage of the agent over the principal.

## 13.6 The Optimization Program of the Principal

According to the timing of the contractual game, the principal must offer a menu of contracts before knowing which type of agent he is facing. Then, the principal's problem writes as

$$\begin{aligned} \max_{\{(\bar{t}, \bar{q}); (\underline{t}, \underline{q})\}} & \nu(S(\underline{q}) - \underline{t}) + (1 - \nu)(S(\bar{q}) - \bar{t}) \\ & \text{subject to (13.9) to (13.12).} \end{aligned}$$

Using the definition of the information rents  $\underline{U} = \underline{t} - \underline{\theta}\underline{q}$  and  $\bar{U} = \bar{t} - \bar{\theta}\bar{q}$ , we can replace transfers in the principal's objective function as functions of information rents and outputs so that the new optimization variables are now  $\{(\underline{U}, \underline{q}); (\bar{U}, \bar{q})\}$ . The focus on information rents enables us to assess the distributive impact of asymmetric information, and the



focus on outputs allows us to analyze its impact on allocative efficiency and the overall gains from trade. Thus an allocation corresponds to a volume of trade and a distribution of the gains from trade between the principal and the agent.

With this change of variables, the principal's objective function can then be rewritten as

$$\underbrace{\nu(S(\underline{q}) - \underline{\theta}q) + (1 - \nu)(S(\bar{q}) - \bar{\theta}\bar{q})}_{\text{expected allocative efficiency}} - \underbrace{(\nu\underline{U} + (1 - \nu)\bar{U})}_{\text{expected information rent}}. \quad (13.20)$$

The first term denotes expected allocative efficiency, and the second term denotes expected information rent which implies that the principal is ready to accept some distortions away from efficiency in order to decrease the agent's information rent.

The incentive constraints (13.9) and (13.10), written in terms of information rents and outputs, becomes respectively

$$\underline{U} \geq \bar{U} + \Delta\theta\bar{q}, \quad (13.21)$$

$$\bar{U} \geq \underline{U} - \Delta\theta\underline{q}. \quad (13.22)$$

The participation constraints (13.11) and (13.12) become respectively

$$\underline{U} \geq 0, \quad (13.23)$$

$$\bar{U} \geq 0. \quad (13.24)$$

The principal wishes to solve problem  $(P)$  below:

$$\max_{\{(\underline{U}, \underline{q}); (\bar{U}, \bar{q})\}} \nu(S(\underline{q}) - \underline{\theta}q) + (1 - \nu)(S(\bar{q}) - \bar{\theta}\bar{q}) - (\nu\underline{U} + (1 - \nu)\bar{U})$$

subject to (13.21) to (13.24).

We index the solution to this problem with a superscript  $SB$ , meaning second-best.

## 13.7 The Rent Extraction-Efficiency Trade-Off

### 13.7.1 The Optimal Contract Under Asymmetric Information

The major technical difficulty of problem  $(P)$  is to determine which of the many constraints imposed by incentive compatibility and participation are the relevant ones. i.e., the binding ones at the optimum or the principal's problem.

Let us first consider contracts without shutdown, i.e., such that  $\bar{q} > 0$ . This is true when the so-called Inada condition  $S'(0) = +\infty$  is satisfied and  $\lim_{q \rightarrow 0} S'(q)q = 0$ .

Note that the  $\underline{\theta}$ -agent's participation constraint (13.23) is always strictly-satisfied. Indeed, (13.24) and (13.21) immediately imply (13.23). (13.22) also seems irrelevant because the difficulty comes from a  $\underline{\theta}$ -agent willing to claim that he is inefficient rather than the reverse.

This simplification in the number of relevant constraints leaves us with only two remaining constraints, the  $\underline{\theta}$ -agent's incentive compatible constraint (13.21) and the  $\bar{\theta}$ -agent's participation constraint (13.24), and both constraints must be binding at the optimum of the principal's problem ( $P$ ):

$$\underline{U} = \Delta\theta\bar{q} \quad (13.25)$$

and

$$\bar{U} = 0. \quad (13.26)$$

Substituting (13.25) and (13.26) into the principal's objective function, we obtain a reduced program ( $P'$ ) with outputs as the only choice variables:

$$\max_{\{\underline{q}, \bar{q}\}} \nu(S(\underline{q}) - \underline{\theta}\underline{q}) + (1 - \nu)(S(\bar{q}) - \bar{\theta}\bar{q}) - (\nu\Delta\theta\bar{q}).$$

Compared with the full information setting, asymmetric information alters the principal's optimization simply by the subtraction of the expected rent that has to be given up to the efficient type. The inefficient type gets no rent, but the efficient type  $\underline{\theta}$  gets information rent that he could obtain by mimicking the inefficient type  $\theta$ . This rent depends only on the level of production requested from this inefficient type.

The first order conditions are then given by

$$S'(\underline{q}^{SB}) = \underline{\theta} \quad \text{or} \quad \underline{q}^{SB} = \underline{q}^*. \quad (13.27)$$

and

$$(1 - \nu)(S'(\bar{q}^{SB}) - \bar{\theta}) = \nu\Delta\theta. \quad (13.28)$$

(13.28) expresses the important trade-off between efficiency and rent extraction which arises under asymmetric information.

To validate our approach based on the sole consideration of the efficient type's incentive compatible constraint, it is necessary to check that the omitted incentive compatible

constraint of an inefficient agent is satisfied. i.e.,  $0 \geq \Delta\theta\bar{q}^{SB} - \Delta\theta\underline{q}^{SB}$ . This latter inequality follows from the monotonicity of the second-best schedule of outputs since we have  $\underline{q}^{SB} = \underline{q}^* > \bar{q}^* > \bar{q}^{SB}$ .

In summary, we have the following proposition.

**Proposition 13.7.1** *Under asymmetric information, the optimal contracts entail:*

- (1) *No output distortion for the efficient type in respect to the first-best,  $\underline{q}^{SB} = \underline{q}^*$ . A downward output distortion for the inefficient type,  $\bar{q}^{SB} < \bar{q}^*$  with*

$$S'(\bar{q}^{SB}) = \bar{\theta} + \frac{\nu}{1-\nu}\Delta\theta. \quad (13.29)$$

- (2) *Only the efficient type gets a positive information rent given by*

$$\underline{U}^{SB} = \Delta\theta\bar{q}^{SB}. \quad (13.30)$$

- (3) *The second-best transfers are respectively given by  $\underline{t}^{SB} = \underline{\theta}\underline{q}^* + \Delta\theta\bar{q}^{SB}$  and  $\bar{t}^{SB} = \bar{\theta}\bar{q}^{SB}$ .*

**Remark 13.7.1** The basic idea and insides of principal-agent theory was somewhat revealed in the book of “the Art of War” written by Sun Tzu in the 6-th century BC. It was the ancient Chines military treatise. The book was China’s earliest and most outstanding and complete work on warcraft. Dubbed ”the Bible of Military Science”, it is also the earliest work on military strategies in the world. He considered the critical importance of information by the saying that: “If you know the enemy and know yourself, you need not fear the result of a hundred battles. If you know yourself but not the enemy, for every victory gained you will also suffer a defeat. If you know neither the enemy nor yourself, you will succumb in every battle.” That is, the best you can do is the first best when information is complete, the best you can do is the second best when information is asymmetric. It would be worse you know thing about the others and yourself.

### 13.7.2 A Graphical Representation of the Second-Best Outcome

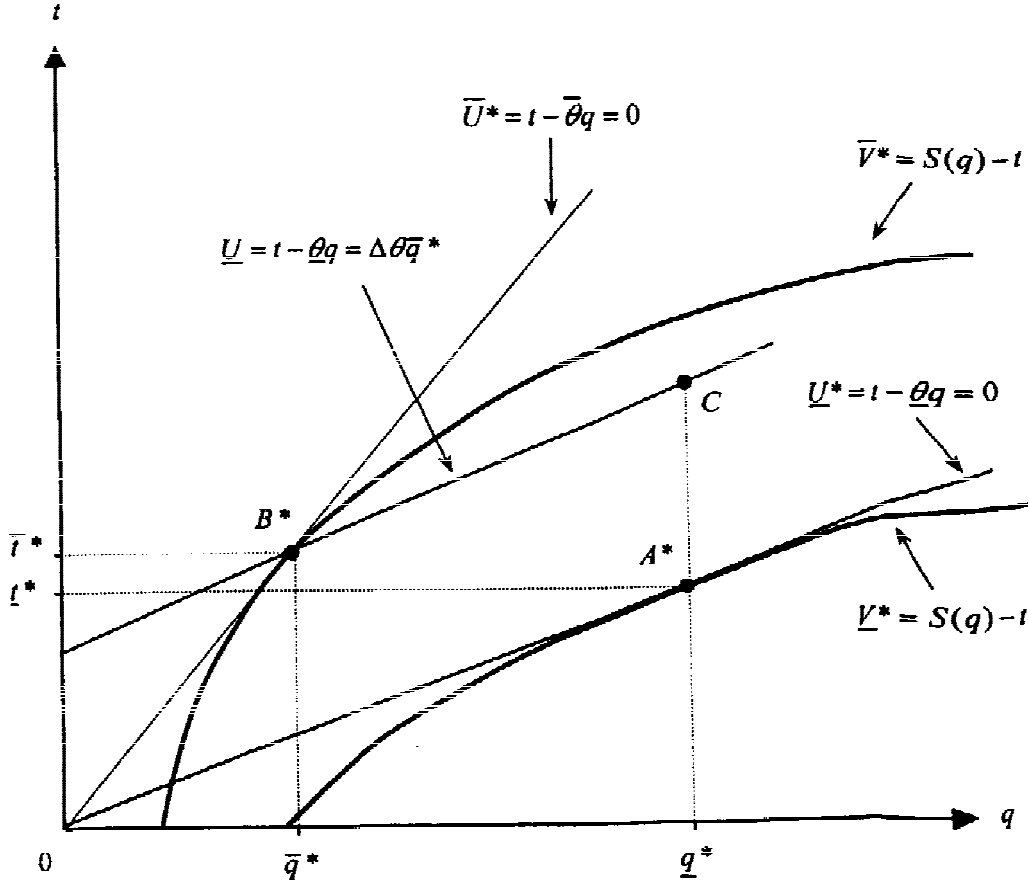


Figure 13.4: Rent needed to implement the first best outputs.

Starting from the complete information optimal contract  $(A^*, B^*)$  that is not incentive compatible, we can construct an incentive compatible contract  $(B^*, C)$  with the same production levels by giving a higher transfer to the agent producing  $q^*$  as shown in the figure above. The contract  $C$  is on the  $\underline{\theta}$ -agent's indifference curve passing through  $B^*$ . Hence, the  $\underline{\theta}$ -agent is now indifferent between  $B^*$  and  $C$ .  $(B^*, C)$  becomes an incentive-compatible menu of contracts. The rent that is given up to the  $\underline{\theta}$ -firm is now  $\Delta\theta\bar{q}^*$ . This contract is not optimal by the first order conditions (13.27) and (13.28). The optimal trade-off finally occurs at  $(A^{SB}, B^{SB})$  as shown in the figure below.

### 13.7.3 Shutdown Policy

If the first-order condition in (13.29) has no positive solution,  $\bar{q}^{SB}$  should be set at zero. We are in the special case of a contract with shutdown.  $B^{SB}$  coincides with 0 and  $A^{SB}$

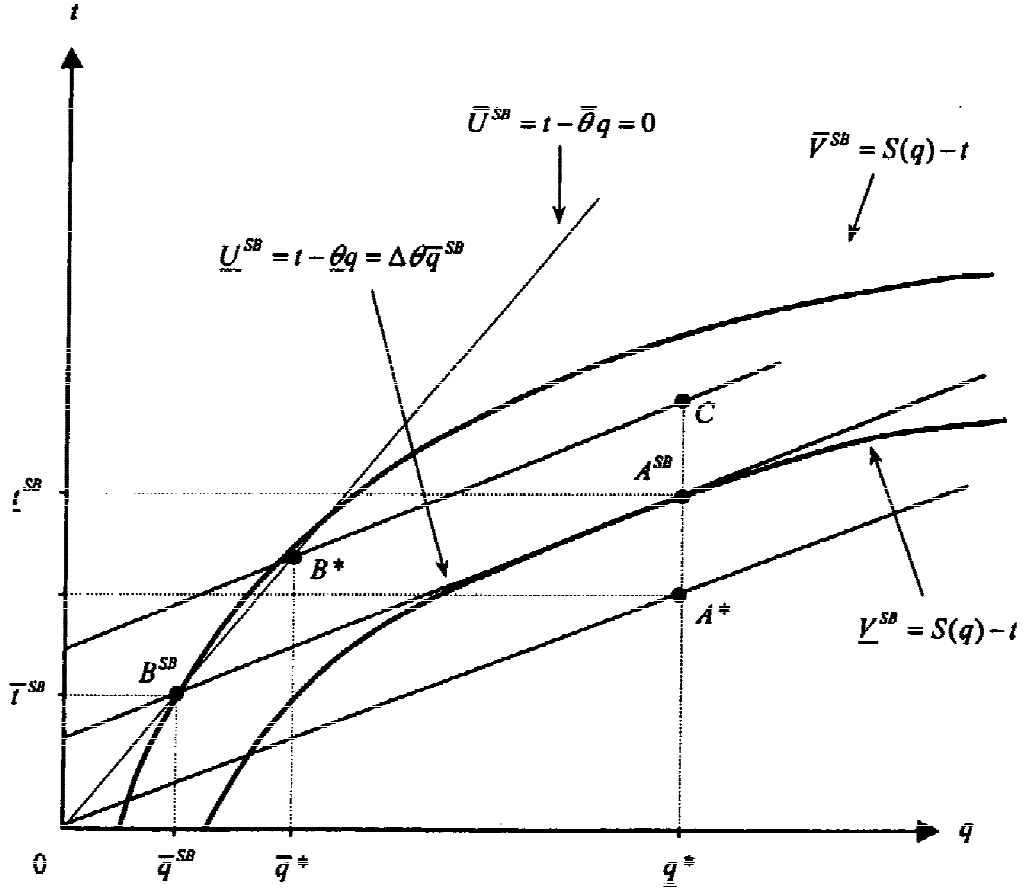


Figure 13.5: Optimal second-best contract  $S^{SB}$  and  $B^{SB}$ .

with  $A^*$  in the figure above. No rent is given up to the  $\underline{\theta}$ -firm by the unique non-null contract  $(\underline{t}^*, \underline{q}^*)$  offered and selected only by agent  $\underline{\theta}$ . The benefit of such a policy is that no rent is given up to the efficient type.

**Remark 13.7.2** The shutdown policy is dependent on the status quo utility levels. Suppose that, for both types, the status quo utility level is  $U_0 > 0$ . Then, from the principal's objective function, we have

$$\frac{\nu}{1-\nu} \Delta\theta \bar{q}^{SB} + U_0 \geq S(\bar{q}^{SB}) - \bar{\theta} \bar{q}^{SB}. \quad (13.31)$$

Thus, for  $\nu$  large enough, shutdown occurs even if the Inada condition  $S'(0) = +\infty$  is satisfied. Note that this case also occurs when the agent has a strictly positive fixed cost  $F > 0$  (to see that, just set  $U_0 = F$ ).

The occurrence of shutdown can also be interpreted as saying that the principal has another choice variable to solve the screening problem. This extra variable is the subset of

types, which are induced to produce a positive amount. Reducing the subset of producing agents obviously reduces the rent of the most efficient type.

## 13.8 The Theory of the Firm Under Asymmetric Information

When the delegation of task occurs within the firm, a major conclusion of the above analysis is that, because of asymmetric information, the firm does not maximize the social value of trade, or more precisely its profit, a maintained assumption of most economic theory. *This lack of allocative efficiency should not be considered as a failure in the rational use of resources within the firm.* Indeed, the point is that allocative efficiency is only one part of the principal's objective. The allocation of resources within the firm remains constrained optimal once informational constraints are fully taken into account.

Williamson (1975) has advanced the view that various transaction costs may impede the achievement of economic transactions. Among the many origins of these costs, Williamson stresses informational impact as an important source of inefficiency. *Even in a world with a costless enforcement of contracts, a major source of allocative inefficiency is the existence of asymmetric information between trading partners.*

*Even though asymmetric information generates allocative inefficiencies, those efficiencies do not call for any public policy motivated by reasons of pure efficiency.* Indeed, any benevolent policymaker in charge of correcting these inefficiencies would face the same informational constraints as the principal. The allocation obtained above is Pareto optimal in the set of incentive feasible allocations or incentive Pareto optimal.

## 13.9 Asymmetric Information and Marginal Cost Pricing

Under complete information, the first-best rules can be interpreted as price equal to marginal cost since consumers on the market will equate their marginal utility of consumption to price.

Under asymmetric information, price equates marginal cost only when the producing

firm is efficient ( $\theta = \underline{\theta}$ ). Using (13.29), we get the expression of the price  $p(\bar{\theta})$  for the inefficient types output

$$p(\bar{\theta}) = \bar{\theta} + \frac{\nu}{1-\nu} \Delta\theta. \quad (13.32)$$

Price is higher than marginal cost in order to decrease the quantity  $\bar{q}$  produced by the inefficient firm and reduce the efficient firm's information rent. Alternatively, we can say that price is equal to a generalized (or virtual) marginal cost that includes, in addition to the traditional marginal cost of the inefficient type  $\bar{\theta}$ , an information cost that is worth  $\frac{\nu}{1-\nu} \Delta\theta$ .

## 13.10 The Revelation Principle

In the above analysis, we have restricted the principal to offer a menu of contracts, one for each possible type. One may wonder if a better outcome could be achieved with a more complex contract allowing the agent possibly to choose among more options. The revelation principle ensures that there is no loss of generality in restricting the principal to offer simple menus having at most as many options as the cardinality of the type space. Those simple menus are actually examples of direct revelation mechanisms.

**Definition 13.10.1** *A direct revelation mechanism is a mapping  $g(\cdot)$  from  $\Theta$  to  $\mathcal{A}$  which writes as  $g(\theta) = (q(\theta), t(\theta))$  for all  $\theta$  belonging to  $\Theta$ . The principal commits to offer the transfer  $t(\tilde{\theta})$  and the production level  $q(\tilde{\theta})$  if the agent announces the value  $\tilde{\theta}$  for any  $\tilde{\theta}$  belonging to  $\Theta$ .*

**Definition 13.10.2** *A direct revelation mechanism  $g(\cdot)$  is truthful if it is incentive compatible for the agent to announce his true type for any type, i.e., if the direct revelation mechanism satisfies the following incentive compatibility constraints:*

$$t(\underline{\theta}) - \underline{\theta}q(\underline{\theta}) \geq t(\bar{\theta}) - \underline{\theta}q(\bar{\theta}), \quad (13.33)$$

$$t(\bar{\theta}) - \bar{\theta}q(\bar{\theta}) \geq t(\underline{\theta}) - \bar{\theta}q(\underline{\theta}). \quad (13.34)$$

Denoting transfer and output for each possible report respectively as  $t(\underline{\theta}) = \underline{t}, q(\underline{\theta}) = \underline{q}$ ,  $t(\bar{\theta}) = \bar{t}$  and  $q(\bar{\theta}) = \bar{q}$ , we get back to the notations of the previous sections.

A more general mechanism can be obtained when communication between the principal and the agent is more complex than simply having the agent report his type to the principal.

Let  $M$  be the message space offered to the agent by a more general mechanism.

**Definition 13.10.3** *A mechanism is a message space  $M$  and a mapping  $\tilde{g}(\cdot)$  from  $M$  to  $\mathcal{A}$  which writes as  $\tilde{g}(m) = (\tilde{q}(m), \tilde{t}(m))$  for all  $m$  belonging to  $M$ .*

When facing such a mechanism, the agent with type  $\theta$  chooses a best message  $m^*(\theta)$  that is implicitly defined as

$$\tilde{t}(m^*(\theta)) - \theta \tilde{q}(m^*(\theta)) \geq \tilde{t}(\tilde{m}) - \theta \tilde{q}(\tilde{m}) \quad \text{for all } \tilde{m} \in M. \quad (13.35)$$

The mechanism  $(M, \tilde{g}(\cdot))$  induces therefore an allocation rule  $a(\theta) = (\tilde{q}(m^*(\theta)), \tilde{t}(m^*(\theta)))$  mapping the set of types  $\Theta$  into the set of allocations  $\mathcal{A}$ .

Then we have the following revelation principle in the one agent case.

**Proposition 13.10.1** *Any allocation rule  $a(\theta)$  obtained with a mechanism  $(M, \tilde{g}(\cdot))$  can also be implemented with a truthful direct revelation mechanism.*

Proof. The indirect mechanism  $(M, \tilde{g}(\cdot))$  induces an allocation rule  $a(\theta) = (\tilde{q}(m^*(\theta)), \tilde{t}(m^*(\theta)))$  from  $M$  into  $\mathcal{A}$ . By composition of  $\tilde{g}(\cdot)$  and  $m^*(\cdot)$ , we can construct a direct revelation mechanism  $g(\cdot)$  mapping  $\Theta$  into  $\mathcal{A}$ , namely  $g = \tilde{g} \circ m^*$ , or more precisely  $g(\theta) = (q(\theta), t(\theta)) \equiv \tilde{g}(m^*(\theta)) = (\tilde{q}(m^*(\theta)), \tilde{t}(m^*(\theta)))$  for all  $\theta \in \Theta$ .

We check now that the direct revelation mechanism  $g(\cdot)$  is truthful. Indeed, since (13.35) is true for all  $\tilde{m}$ , it holds in particular for  $\tilde{m} = m^*(\theta')$  for all  $\theta' \in \Theta$ . Thus we have

$$\tilde{t}(m^*(\theta)) - \theta \tilde{q}(m^*(\theta)) \geq \tilde{t}(m^*(\theta')) - \theta \tilde{q}(m^*(\theta')) \quad \text{for all } (\theta, \theta') \in \Theta^2. \quad (13.36)$$

Finally, using the definition of  $g(\cdot)$ , we get

$$t(\theta) - \theta q(\theta) \geq t(\theta') - \theta q(\theta') \quad \text{for all } (\theta, \theta') \in \Theta^2. \quad (13.37)$$

Hence, the direct revelation mechanism  $g(\cdot)$  is truthful.

Importantly, the revelation principle provides a considerable simplification of contract theory. It enables us to restrict the analysis to a simple and well-defined family of functions, the truthful direct revelation mechanism.



## 13.11 A More General Utility Function for the Agent

Still keeping quasi-linear utility functions, let  $U = t - C(q, \theta)$  now be the agent's objective function in the assumptions:  $C_q > 0$ ,  $C_\theta > 0$ ,  $C_{qq} > 0$  and  $C_{q\theta} > 0$ . The generalization of the Spence-Mirrlees property is now  $C_{q\theta} > 0$ . This latter condition still ensures that the different types of the agent have indifference curves which cross each other at most once. This Spence-Mirrlees property is quite clear: a more efficient type is also more efficient at the margin.

Incentive feasible allocations satisfy the following incentive and participation constraints:

$$\underline{U} = \underline{t} - C(\underline{q}, \underline{\theta}) \geq \bar{t} - C(\bar{q}, \underline{\theta}), \quad (13.38)$$

$$\bar{U} = \bar{t} - C(\bar{q}, \bar{\theta}) \geq \underline{t} - C(\underline{q}, \bar{\theta}), \quad (13.39)$$

$$\underline{U} = \underline{t} - C(\underline{q}, \underline{\theta}) \geq 0, \quad (13.40)$$

$$\bar{U} = \bar{t} - C(\bar{q}, \bar{\theta}) \geq 0. \quad (13.41)$$

### 13.11.1 The Optimal Contract

Just as before, the incentive constraint of an efficient type in (13.38) and the participation constraint of an inefficient type in (13.41) are the two relevant constraints for optimization. These constraints rewrite respectively as

$$\underline{U} \geq \bar{U} + \Phi(\bar{q}) \quad (13.42)$$

where  $\Phi(\bar{q}) = C(\bar{q}, \bar{\theta}) - C(\bar{q}, \underline{\theta})$  (with  $\Phi' > 0$  and  $\Phi'' > 0$ ), and

$$\bar{U} \geq 0. \quad (13.43)$$

Those constraints are both binding at the second-best optimum, which leads to the following expression of the efficient type's rent

$$\underline{U} = \Phi(\bar{q}). \quad (13.44)$$

Since  $\Phi' > 0$ , reducing the inefficient agent's output also reduces the efficient agent's information rent.

With the assumptions made on  $C(\cdot)$ , one can also check that the principal's objective function is strictly concave with respect to outputs.

The solution of the principal's program can be summarized as follows:

**Proposition 13.11.1** *With general preferences satisfying the Spence-Mirrlees property,  $C_{q\theta} > 0$ , the optimal menu of contracts entails:*

- (1) *No output distortion with respect to the first-best outcome for the efficient type,  $q^{SB} = \underline{q}^*$  with*

$$S'(\underline{q}^*) = C_q(\underline{q}^*, \underline{\theta}). \quad (13.45)$$

*A downward output distortion for the inefficient type,  $\bar{q}^{SB} < \bar{q}^*$  with*

$$S'(\bar{q}^*) = C_q(\bar{q}^*, \bar{\theta}) \quad (13.46)$$

*and*

$$S'(\bar{q}^{SB}) = C_q(\bar{q}^{SB}, \bar{\theta}) + \frac{\nu}{1-\nu} \Phi'(\bar{q}^{SB}). \quad (13.47)$$

- (2) *Only the efficient type gets a positive information rent given by  $\underline{U}^{SB} = \Phi(\bar{q}^{SB})$ .*

- (3) *The second-best transfers are respectively given by  $\underline{t}^{SB} = C(\underline{q}^*, \underline{\theta}) + \Phi(\bar{q}^{SB})$  and  $\bar{t}^{SB} = C(\bar{q}^{SB}, \bar{\theta})$ .*

The first-order conditions (13.45) and (13.47) characterize the optimal solution if the neglected incentive constraint (13.39) is satisfied. For this to be true, we need to have

$$\begin{aligned} \bar{t}^{SB} - C(\bar{q}^{SB}, \bar{\theta}) &\geq \underline{t}^{SB} - C(\underline{q}^{SB}, \bar{\theta}), \\ &= \bar{t}^{SB} - C(\bar{q}^{SB}, \underline{\theta}) + C(\underline{q}^{SB}, \underline{\theta}) - C(\underline{q}^{SB}, \bar{\theta}) \end{aligned} \quad (13.48)$$

by noting that (13.38) holds with equality at the optimal output such that  $\underline{t}^{SB} = \bar{t}^{SB} - C(\bar{q}^{SB}, \underline{\theta}) + C(\underline{q}^{SB}, \underline{\theta})$ . Thus, we need to have

$$0 \geq \Phi(\bar{q}^{SB}) - \Phi(\underline{q}^{SB}). \quad (13.49)$$

Since  $\Phi' > 0$  from the Spence-Mirrlees property, then (13.49) is equivalent to  $\bar{q}^{SB} \leq \underline{q}^{SB}$ . But from our assumptions we easily derive that  $\underline{q}^{SB} = \underline{q}^* > \bar{q}^* > \bar{q}^{SB}$ . So the Spence-Mirrlees property guarantees that only the efficient type's incentive compatible constraint has to be taken into account.

**Remark 13.11.1** When there are the joint presence of asymmetric information and network externalities, “no distortion on the top” rule may not be true so that the most efficient agent may not have the first-best outcome. Meng and Tian (2008) showed that the optimal contract for the principal may exhibit two-way distortion: the output of any agent is oversupplied (relative to the first-best) when his marginal cost of effort is low, and undersupplied when his marginal cost of effort is high.

The “network externalities” might arise for any of the following reasons: because the usefulness of the commodity depends directly on the size of the network (e.g., telephones, fax machines); or because of bandwagon effect, which means the desire to be in style: to have a good because almost everyone else has it; or indirectly through the availability of complementary goods and services (often known as the “hardware-software paradigm” ) or of postpurchase services (e.g., for automobiles). Although “network externalities” are often regarded as positive impact on each others’ consumption, it may display negative properties in some cases. For example, people has the desire to own exclusive or unique goods, which is called “Snob effect”. The quantity demanded of a “snob” good is higher the fewer the people who own it.

### 13.11.2 More than One Good

Let us now assume that the agent is producing a whole vector of goods  $q = (q_1, \dots, q_n)$  for the principal. The agents’ cost function becomes  $C(q, \theta)$  with  $C(\cdot)$  being strictly convex in  $q$ . The value for the principal of consuming this whole bundle is now  $S(q)$  with  $S(\cdot)$  being strictly concave in  $q$ .

In this multi-output incentive problem, the principal is interested in a whole set of activities carried out simultaneously by the agent. It is straightforward to check that the efficient agent’s information rent is now written as  $\underline{U} = \Phi(q)$  with  $\Phi(q) = C(q, \bar{\theta}) - C(q, \underline{\theta})$ . This leads to second-best optimal outputs. The efficient type produces the first-best vector of outputs  $\underline{q}^{SB} = \underline{q}^*$  with

$$S_{q_i}(\underline{q}^*) = C_{q_i}(\underline{q}^*, \underline{\theta}) \quad \text{for all } i \in \{1, \dots, n\}. \quad (13.50)$$

The inefficient types vector of outputs  $\bar{q}^{SB}$  is instead characterized by the first-order

conditions

$$S_{q_i}(\bar{q}^{SB}) = C_{q_i}(\bar{q}^{SB}, \bar{\theta}) + \frac{\nu}{1-\nu} \Phi_{q_i}(\bar{q}^{SB}) \quad \text{for all } i \in \{1, \dots, n\}, \quad (13.51)$$

which generalizes the distortion of models with a single good.

Without further specifying the value and cost functions, the second-best outputs define a vector of outputs with some components  $\bar{q}_i^{SB}$  above  $\bar{q}_i^*$  for a subset of indices  $i$ .

Turning to incentive compatibility, summing the incentive constraints  $\underline{U} \geq \bar{U} + \Phi(\bar{q})$  and  $\bar{U} \geq \underline{U} - \Phi(\underline{q})$  for any incentive feasible contract yields

$$\Phi(\underline{q}) = C(\underline{q}, \bar{\theta}) - C(\underline{q}, \underline{\theta}) \quad (13.52)$$

$$\begin{aligned} &\geq C(\bar{q}, \bar{\theta}) - C(\bar{q}, \underline{\theta}) \\ &= \Phi(\bar{q}) \quad \text{for all implementable pairs } (\bar{q}, \underline{q}). \end{aligned} \quad (13.53)$$

Obviously, this condition is satisfied if the Spence-Mirrlees property  $C_{q_i\theta} > 0$  holds for each output  $i$  and if the monotonicity conditions  $\bar{q}_i < \underline{q}_i$  for all  $i$  are satisfied.

## 13.12 Ex Ante versus Ex Post Participation Constraints

The case of contracts we consider so far is offered at the interim stage, i.e., the agent already knows his type. However, sometimes the principal and the agent can contract at the ex ante stage, i.e., before the agent discovers his type. For instance, the contracts of the firm may be designed before the agent receives any piece of information on his productivity. In this section, we characterize the optimal contract for this alternative timing under various assumptions about the risk aversion of the two players.

### 13.12.1 Risk Neutrality

Suppose that the principal and the agent meet and contract ex ante. If the agent is risk neutral, his ex ante participation constraint is now written as

$$\nu \underline{U} + (1-\nu) \bar{U} \geq 0. \quad (13.54)$$

This ex ante participation constraint replaces the two interim participation constraints.

Since the principal's objective function is decreasing in the agent's expected information rent, the principal wants to impose a zero expected rent to the agent and have (13.54) be binding. Moreover, the principal must structure the rents  $\underline{U}$  and  $\bar{U}$  to ensure that the two incentive constraints remain satisfied. An example of such a rent distribution that is both incentive compatible and satisfies the ex ante participation constraint with an equality is

$$\underline{U}^* = (1 - \nu)\Delta\theta\bar{q}^* > 0 \quad \text{and} \quad \bar{U}^* = -\nu\Delta\theta\bar{q}^* < 0. \quad (13.55)$$

With such a rent distribution, the optimal contract implements the first-best outputs without cost from the principal's point of view as long as the first-best is monotonic as requested by the implementability condition. In the contract defined by (13.55), the agent is rewarded when he is efficient and punished when he turns out to be inefficient. In summary, we have

**Proposition 13.12.1** *When the agent is risk neutral and contracting takes place ex ante, the optimal incentive contract implements the first-best outcome.*

**Remark 13.12.1** The principal has in fact more options in structuring the rents  $\underline{U}$  and  $\bar{U}$  in such a way that the incentive compatible constraints hold and the ex ante participation constraint (13.54) holds with an equality. Consider the following contracts  $\{(\underline{t}^*, \underline{q}^*); (\bar{t}^*, \bar{q}^*)\}$  where  $\underline{t}^* = S(\underline{q}^*) - T^*$  and  $\bar{t}^* = S(\bar{q}^*) - T^*$ , with  $T^*$  being a lump-sum payment to be defined below. This contract is incentive compatible since

$$\underline{t}^* - \theta\underline{q}^* = S(\underline{q}^*) - \theta\underline{q}^* - T^* > S(\bar{q}^*) - \theta\bar{q}^* - T^* = \bar{t}^* - \theta\bar{q}^* \quad (13.56)$$

by definition of  $\underline{q}^*$ , and

$$\bar{t}^* - \bar{\theta}\bar{q}^* = S(\bar{q}^*) - \bar{\theta}\bar{q}^* - T^* > S(\underline{q}^*) - \bar{\theta}\underline{q}^* - T^* = \underline{t}^* - \bar{\theta}\underline{q}^* \quad (13.57)$$

by definition of  $\bar{q}^*$ .

Note that the incentive compatibility constraints are now strict inequalities. Moreover, the fixed-fee  $T^*$  can be used to satisfy the agent's ex ante participation constraint with an equality by choosing  $T^* = \nu(S(\underline{q}^*) - \theta\underline{q}^*) + (1 - \nu)(S(\bar{q}^*) - \bar{\theta}\bar{q}^*)$ . This implementation of the first-best outcome amounts to having the principal selling the benefit of the relationship to the risk-neutral agent for a fixed up-front payment  $T^*$ . The agent benefits from the

full value of the good and trades off the value of any production against its cost just as if he was an efficiency maximizer. We will say that the agent is residual claimant for the firms profit.

### 13.12.2 Risk Aversion

#### A Risk-Averse Agent

The previous section has shown us that the implementation of the first-best is feasible with risk neutrality. What happens if the agent is risk-averse?

Consider now a risk-averse agent with a Von Neumann-Morgenstern utility function  $u(\cdot)$  defined on his monetary gains  $t - \theta q$ , such that  $u' > 0$ ,  $u'' < 0$  and  $u(0) = 0$ . Again, the contract between the principal and the agent is signed before the agent discovers his type. The incentive-compatibility constraints are unchanged but the agent's ex ante participation constraint is now written as

$$\nu u(\underline{U}) + (1 - \nu)u(\bar{U}) \geq 0. \quad (13.58)$$

As usual, one can check incentive-compatibility constraint (13.22) for the inefficient agent is slack (not binding) at the optimum, and thus the principal's program reduces now to

$$\max_{\{(\bar{U}, \bar{q}); (\underline{U}, \underline{q})\}} \nu(S(\underline{q}) - \underline{\theta}\underline{q} - \underline{U}) + (1 - \nu)(S(\bar{q}) - \bar{\theta}\bar{q} - \bar{U}),$$

subject to (13.21) and (13.58).

We have the following proposition.

**Proposition 13.12.2** *When the agent is risk-averse and contracting takes place ex ante, the optimal menu of contracts entails:*

- (1) *No output distortion for the efficient  $\underline{q}^{SB} = \underline{q}^*$ . A downward output distortion for the inefficient type  $\bar{q}^{SB} < \bar{q}^*$ , with*

$$S'(\bar{q}^{SB}) = \bar{\theta} + \frac{\nu(u'(\bar{U}^{SB}) - u'(\underline{U}^{SB}))}{\nu u'(\underline{U}^{SB}) + (1 - \nu)u'(\bar{U}^{SB})} \Delta\theta. \quad (13.59)$$

- (2) *Both (13.21) and (13.58) are the only binding constraints. The efficient (resp. inefficient) type gets a strictly positive (resp. negative) ex post information rent,  $\underline{U}^{SB} > 0 > \bar{U}^{SB}$ .*

Proof: Define the following Lagrangian for the principals problem

$$\begin{aligned} L(\underline{q}, \bar{q}, \underline{U}, \bar{U}, \lambda, \mu) = & \nu(S(\underline{q}) - \underline{\theta}\underline{q} - \underline{U}) + (1 - \nu)(S(\bar{q}) - \bar{\theta}\bar{q} - \bar{U}) \\ & + \lambda(\underline{U} - \bar{U} - \Delta\theta\bar{q}) + \mu(\nu u(\underline{U}) + (1 - \nu)u(\bar{U})). \end{aligned} \quad (13.60)$$

Optimizing w.r.t.  $\underline{U}$  and  $\bar{U}$  yields respectively

$$-\nu + \lambda + \mu\nu u'(\underline{U}^{SB}) = 0 \quad (13.61)$$

$$-(1 - \nu) - \lambda + \mu(1 - \nu)u'(\bar{U}^{SB}) = 0. \quad (13.62)$$

Summing the above two equations, we obtain

$$\mu(\nu u'(\underline{U}^{SB}) + (1 - \nu)u'(\bar{U}^{SB})) = 1. \quad (13.63)$$

and thus  $\mu > 0$ . Using (13.63) and inserting it into (13.61) yields

$$\lambda = \frac{\nu(1 - \nu)(u'(\bar{U}^{SB}) - u'(\underline{U}^{SB}))}{\nu u'(\underline{U}^{SB}) + (1 - \nu)u'(\bar{U}^{SB})}. \quad (13.64)$$

Moreover, (13.21) implies that  $\underline{U}^{SB} \geq \bar{U}^{SB}$  and thus  $\lambda \geq 0$ , with  $\lambda > 0$  for a positive output  $y$ .

Optimizing with respect to outputs yields respectively

$$S'(\underline{q}^{SB}) = \underline{\theta} \quad (13.65)$$

and

$$S'(\bar{q}^{SB}) = \bar{\theta} + \frac{\lambda}{1 - \nu}\Delta\theta. \quad (13.66)$$

Simplifying by using (13.64) yields (13.59).

Thus, with risk aversion, the principal can no longer costlessly structure the agent's information rents to ensure the efficient type's incentive compatibility constraint. Creating a wedge between  $\underline{U}$  and  $\bar{U}$  to satisfy (13.21) makes the risk-averse agent bear some risk. To guarantee the participation of the risk-averse agent, the principal must now pay a risk premium. Reducing this premium calls for a downward reduction in the inefficient type's output so that the risk borne by the agent is lower. As expected, the agent's risk aversion leads the principal to weaken the incentives.

When the agent becomes infinitely risk averse, everything happens as if he had an ex post individual rationality constraint for the worst state of the world given by (13.24).

In the limit, the inefficient agent's output  $\bar{q}^{SB}$  and the utility levels  $\underline{U}^{SB}$  and  $\bar{U}^{SB}$  all converge toward the same solution. So, the previous model at the interim stage can also be interpreted as a model with an ex ante infinitely risk-agent at the zero utility level.

### A Risk-Averse Principal

Consider now a risk-averse principal with a Von Neumann-Morgenstern utility function  $v(\cdot)$  defined on his monetary gains from trade  $S(q) - t$  such that  $v' > 0$ ,  $v'' < 0$  and  $v(0) = 0$ . Again, the contract between the principal and the risk-neutral agent is signed before the agent knows his type.

In this context, the first-best contract obviously calls for the first-best output  $\underline{q}^*$  and  $\bar{q}^*$  being produced. It also calls for the principal to be fully insured between both states of nature and for the agent's ex ante participation constraint to be binding. This leads us to the following two conditions that must be satisfied by the agent's rents  $\underline{U}^*$  and  $\bar{U}^*$ :

$$S(\underline{q}^*) - \underline{\theta}\underline{q}^* - \underline{U}^* = S(\bar{q}^*) - \bar{\theta}\bar{q}^* - \bar{U}^* \quad (13.67)$$

and

$$\nu\underline{U}^* + (1 - \nu)\bar{U}^* = 0. \quad (13.68)$$

Solving this system of two equations with two unknowns  $(\underline{U}^*, \bar{U}^*)$  yields

$$\underline{U}^* = (1 - \nu)(S(\underline{q}^*) - \underline{\theta}\underline{q}^* - (S(\bar{q}^*) - \bar{\theta}\bar{q}^*)) \quad (13.69)$$

and

$$\bar{U}^* = -\nu(S(\underline{q}^*) - \underline{\theta}\underline{q}^* - (S(\bar{q}^*) - \bar{\theta}\bar{q}^*)). \quad (13.70)$$

Note that the first-best profile of information rents satisfies both types' incentive compatibility constraints since

$$\underline{U}^* - \bar{U}^* = S(\underline{q}^*) - \underline{\theta}\underline{q}^* - (S(\bar{q}^*) - \bar{\theta}\bar{q}^*) > \Delta\theta\bar{q}^* \quad (13.71)$$

(from the definition of  $\underline{q}^*$ ) and

$$\bar{U}^* - \underline{U}^* = S(\bar{q}^*) - \bar{\theta}\bar{q}^* - (S(\underline{q}^*) - \underline{\theta}\underline{q}^*) > -\Delta\theta\underline{q}^*, \quad (13.72)$$

(from the definition of  $\bar{q}^*$ ). Hence, the profile of rents  $(\underline{U}^*, \bar{U}^*)$  is incentive compatible and the first-best allocation is easily implemented in this framework. We can thus generalize the proposition for the case of risk neutral as follows:



**Proposition 13.12.3** *When the principal is risk-averse over the monetary gains  $S(q) - t$ , the agent is risk-neutral, and contracting takes place *ex ante*, the optimal incentive contract implements the first-best outcome.*

**Remark 13.12.2** It is interesting to note that  $\underline{U}^*$  and  $\bar{U}^*$  obtained in (13.69) and (13.70) are also the levels of rent obtained in (13.56) and (13.57). Indeed, the lump-sum payment  $T^* = \nu(S(\underline{q}^*) - \underline{\theta}\underline{q}^*) + (1 - \nu)(S(\bar{q}^*) - \bar{\theta}\bar{q}^*)$ , which allows the principal to make the risk-neutral agent residual claimant for the hierarchy's profit, also provides full insurance to the principal. By making the risk-neutral agent the residual claimant for the value of trade, *ex ante* contracting allows the risk-averse principal to get full insurance and implement the first-best outcome despite the informational problem.

Of course this result does not hold anymore if the agent's interim participation constraints must be satisfied. In this case, we still guess a solution such that (13.23) is slack at the optimum. The principal's program now reduces to:

$$\max_{\{(\bar{U}, \bar{q}); (\underline{U}, \underline{q})\}} \nu v(S(\underline{q}) - \underline{\theta}\underline{q} - \underline{U}) + (1 - \nu)v(S(\bar{q}) - \bar{\theta}\bar{q} - \bar{U})$$

subject to (13.21) to (13.24).

Inserting the values of  $\underline{U}$  and  $\bar{U}$  that were obtained from the binding constraints in (13.21) and (13.24) into the principal's objective function and optimizing with respect to outputs leads to  $\underline{q}^{SB} = \underline{q}^*$ , i.e., no distortion for the efficient type, just as in the ease of risk neutrality and a downward distortion of the inefficient type's output  $\bar{q}^{SB} < \bar{q}^*$  given by

$$S'(\bar{q}^{SB}) = \bar{\theta} + \frac{\nu v'(\underline{V}^{SB})}{(1 - \nu)v'(\bar{V}^{SB})} \Delta\theta. \quad (13.73)$$

where  $\underline{V}^{SB} = S(\underline{q}^*) - \underline{\theta}\underline{q}^* - \Delta\theta\bar{q}^{SB}$  and  $\bar{V}^{SB} = S(\bar{q}^{SB}) - \bar{\theta}\bar{q}^{SB}$  are the principal's payoffs in both states of nature. We can check that  $\bar{V}^{SB} < \underline{V}^{SB}$  since  $S(\bar{q}^{SB}) - \bar{\theta}\bar{q}^{SB} < S(\underline{q}^*) - \underline{\theta}\underline{q}^*$  from the definition of  $\underline{q}^*$ . In particular, we observe that the distortion in the right-hand side of (13.73) is always lower than  $\frac{\nu}{1-\nu}\Delta\theta$ , its value with a risk-neutral principal. The intuition is straightforward. By increasing  $\bar{q}$  above its value with risk neutrality, the risk-averse principal reduces the difference between  $\underline{V}^{SB}$  and  $\bar{V}^{SB}$ . This gives the principal some insurance and increases his *ex ante* payoff.

For example, if  $\nu(x) = \frac{1-e^{-rx}}{r}$ , (13.73) becomes  $S'(\bar{q}^{SB}) = \bar{\theta} + \frac{\nu}{1-\nu}e^{r(\bar{V}^{SB}-\underline{V}^{SB})}\Delta\theta$ . If  $r = 0$ , we get back the distortion obtained before for the case of with a risk-neutral

principal and interim participation constraints for the agent. Since  $\bar{V}^{SB} < \underline{V}^{SB}$ , we observe that the first-best is implemented when  $r$  goes to infinity. In the limit, the infinitely risk-averse principal is only interested in the inefficient state of nature for which he wants to maximize the surplus, since there is no rent for the inefficient agent. Moreover, giving a rent to the efficient agent is now without cost for the principal.

Risk aversion on the side of the principal is quite natural in some contexts. A local regulator with a limited budget or a specialized bank dealing with relatively correlated projects may be insufficiently diversified to become completely risk neutral. See Lewis and Sappington (Rand J. Econ, 1995) for an application to the regulation of public utilities.

## 13.13 Commitment

To solve the incentive problem, we have implicitly assumed that the principal has a strong ability to commit himself not only to a distribution of rents that will induce information revelation but also to some allocative inefficiency designed to reduce the cost of this revelation. Alternatively, this assumption also means that the court of law can perfectly enforce the contract and that neither renegotiating nor reneging on the contract is a feasible alternative for the agent and (or) the principal. What can happen when either of those two assumptions is relaxed?

### 13.13.1 Renegotiating a Contract

Renegotiation is a voluntary act that should benefit both the principal and the agent. It should be contrasted with a breach of contract, which can hurt one of the contracting parties. One should view a renegotiation procedure as the ability of the contracting partners to achieve a Pareto improving trade if any becomes incentive feasible along the course of actions.

Once the different types have revealed themselves to the principal by selecting the contracts  $(\underline{t}^{SB}, \underline{q}^{SB})$  for the efficient type and  $(\bar{t}^{SB}, \bar{q}^{SB})$  for the inefficient type, the principal may propose a renegotiation to get around the allocative inefficiency he has imposed on the inefficient agent's output. The gain from this renegotiation comes from raising allocative efficiency for the inefficient type and moving output from  $\bar{q}^{SB}$  to  $\bar{q}^*$ . To share these

new gains from trade with the inefficient agent, the principal must at least offer him the same utility level as before renegotiation. The participation constraint of the inefficient agent can still be kept at zero when the transfer of this type is raised from  $\bar{t}^{SB} = \bar{\theta}\bar{q}^{SB}$  to  $\bar{t}^* = \bar{\theta}\bar{q}^*$ . However, raising this transfer also hardens the ex ante incentive compatibility constraint of the efficient type. Indeed, it becomes more valuable for an efficient type to hide his type so that he can obtain this larger transfer, and truthful revelation by the efficient type is no longer obtained in equilibrium. There is a fundamental trade-off between raising efficiency ex post and hardening ex ante incentives when renegotiation is an issue.

### 13.13.2 Reneging on a Contract

A second source of imperfection arises when either the principal or the agent reneges on their previous contractual obligation. Let us take the case of the principal reneging on the contract. Indeed, once the agent has revealed himself to the principal by selecting the contract within the menu offered by the principal, the latter, having learned the agent's type, might propose the complete information contract which extracts all rents without inducing inefficiency. On the other hand, the agent may want to renege on a contract which gives him a negative ex post utility level as we discussed before. In this case, the threat of the agent reneging a contract signed at the ex ante stage forces the agent's participation constraints to be written in interim terms. Such a setting justifies the focus on the case of interim contracting.

## 13.14 Informative Signals to Improve Contracting

In this section, we investigate the impacts of various improvements of the principal's information system on the optimal contract. The idea here is to see how signals that are exogenous to the relationship can be used by the principal to better design the contract with the agent.

### 13.14.1 Ex Post Verifiable Signal

Suppose that the principal, the agent and the court of law observe ex post a viable signal  $\sigma$  which is correlated with  $\theta$ . This signal is observed after the agent's choice of production. The contract can then be conditioned on both the agent's report and the observed signal that provides useful information on the underlying state of nature.

For simplicity, assume that this signal may take only two values,  $\sigma_1$  and  $\sigma_2$ . Let the conditional probabilities of these respective realizations of the signal be  $\mu_1 = \Pr(\sigma = \sigma_1/\theta = \underline{\theta}) \geq 1/2$  and  $\mu_2 = \Pr(\sigma = \sigma_2/\theta = \bar{\theta}) \geq 1/2$ . Note that, if  $\mu_1 = \mu_2 = 1/2$ , the signal  $\sigma$  is uninformative. Otherwise,  $\sigma_1$  brings good news the fact that the agent is efficient and  $\sigma_2$  brings bad news, since it is more likely that the agent is inefficient in this case.

Let us adopt the following notations for the ex post information rents:  $u_{11} = t(\underline{\theta}, \sigma_1) - \underline{\theta}q(\underline{\theta}, \sigma_1)$ ,  $u_{12} = t(\underline{\theta}, \sigma_2) - \underline{\theta}q(\underline{\theta}, \sigma_2)$ ,  $u_{21} = t(\bar{\theta}, \sigma_1) - \bar{\theta}q(\bar{\theta}, \sigma_1)$ , and  $u_{22} = t(\bar{\theta}, \sigma_2) - \bar{\theta}q(\bar{\theta}, \sigma_2)$ . Similar notations are used for the outputs  $q_{jj}$ . The agent discovers his type and plays the mechanism before the signal  $\sigma$  realizes. Then the incentive and participation constraints must be written in expectation over the realization of  $\sigma$ . Incentive constraints for both types write respectively as

$$\mu_1 u_{11} + (1 - \mu_1) u_{12} \geq \mu_1 (u_{21} + \Delta\theta q_{21}) + (1 - \mu_1) (u_{22} + \Delta\theta q_{22}) \quad (13.74)$$

$$(1 - \mu_2) u_{21} + \mu_2 u_{22} \geq (1 - \mu_2) (u_{11} - \Delta\theta q_{11}) + \mu_2 (u_{12} - \Delta\theta q_{12}). \quad (13.75)$$

Participation constraints for both types are written as

$$\mu_1 u_{11} + (1 - \mu_1) u_{12} \geq 0, \quad (13.76)$$

$$(1 - \mu_2) u_{21} + \mu_2 u_{22} \geq 0. \quad (13.77)$$

Note that, for a given schedule of output  $q_{ij}$ , the system (13.74) through (13.77) has as many equations as unknowns  $u_{ij}$ . When the determinant of the coefficient matrix of the system (13.74) to (13.77) is nonzero, one can find ex post rents  $u_{ij}$  (or equivalent transfers) such that all these constraints are binding. In this case, the agent receives no rent whatever his type. Moreover, any choice of production levels, in particular the complete information optimal ones, can be implemented this way. Note that the determinant of the system is nonzero when

$$1 - \mu_1 - \mu_2 \neq 0 \quad (13.78)$$

that fails only if  $\mu_1 = \mu_2 = \frac{1}{2}$ , which corresponds to the case of an uninformative and useless signal.

### 13.14.2 Ex Ante Nonverifiable Signal

Now suppose that a nonverifiable binary signal  $\sigma$  about  $\theta$  is available to the principal at the ex ante stage. Before offering an incentive contract, the principal computes, using the Bayes law, his posterior belief that the agent is efficient for each value of this signal, namely

$$\hat{\nu}_1 = Pr(\theta = \underline{\theta}/\sigma = \sigma_1) = \frac{\nu\mu_1}{\nu\mu_1 + (1-\nu)(1-\mu_2)}, \quad (13.79)$$

$$\hat{\nu}_2 = Pr(\theta = \underline{\theta}/\sigma = \sigma_2) = \frac{\nu(1-\mu_1)}{\nu(1-\mu_1) + (1-\nu)\mu_2}. \quad (13.80)$$

Then the optimal contract entails a downward distortion of the inefficient agents production  $\bar{q}^{SB}(\sigma_i)$  which is for signals  $\sigma_1$ , and  $\sigma_2$  respectively:

$$S'(\bar{q}^{SB}(\sigma_1)) = \bar{\theta} + \frac{\hat{\nu}_1}{1-\hat{\nu}_1} \Delta\theta = \bar{\theta} + \frac{\nu\mu_1}{(1-\nu)(1-\mu_2)} \Delta\theta \quad (13.81)$$

$$S'(\bar{q}^{SB}(\sigma_2)) = \bar{\theta} + \frac{\hat{\nu}_2}{1-\hat{\nu}_2} \Delta\theta = \bar{\theta} + \frac{\nu(1-\mu_1)}{(1-\nu)\mu_2} \Delta\theta. \quad (13.82)$$

In the case where  $\mu_1 = \mu_2 = \mu > \frac{1}{2}$ , we can interpret  $\mu$  as an index of the informativeness of the signal. Observing  $\sigma_1$ , the principal thinks that it is more likely that the agent is efficient. A stronger reduction in  $\bar{q}^{SB}$  and thus in the efficient type's information rent is called for after  $\sigma_1$ . (13.81) shows that incentives decrease with respect to the case without informative signal since  $\left(\frac{\mu}{1-\mu} > 1\right)$ . In particular, if  $\mu$  is large enough, the principal shuts down the inefficient firm after having observed  $\sigma_1$ . The principal offers a high-powered incentive contract only to the efficient agent, which leaves him with no rent. On the contrary, because he is less likely to face an efficient type after having observed  $\sigma_2$ , the principal reduces less of the information rent than in the case without an informative signal since  $\left(\frac{1-\mu}{\mu} < 1\right)$ . Incentives are stronger.

## 13.15 Contract Theory at Work

This section proposes several classical settings where the basic model of this chapter is useful. Introducing adverse selection in each of these contexts has proved to be a significant improvement of standard microeconomic analysis.

### 13.15.1 Regulation

In the Baron and Myerson (Econometrica, 1982) regulation model, the principal is a regulator who maximizes a weighted average of the agents' surplus  $S(q) - t$  and of a regulated monopoly's profit  $U = t - \theta q$ , with a weight  $\alpha < 1$  for the firm's profit. The principal's objective function is written now as  $V = S(q) - \theta q - (1 - \alpha)U$ . Because  $\alpha < 1$ , it is socially costly to give up a rent to the firm. Maximizing expected social welfare under incentive and participation constraints leads to  $\bar{q}^{SB} = \underline{q}^*$  for the efficient type and a downward distortion for the inefficient type,  $\bar{q}^{SB} < \bar{q}^*$  which is given by

$$S'(\bar{q}^{SB}) = \bar{\theta} + \frac{\nu}{1 - \nu}(1 - \alpha)\Delta\theta. \quad (13.83)$$

Note that a higher value of  $\alpha$  reduces the output distortion, because the regulator is less concerned by the distribution of rents within society as  $\alpha$  increases. If  $\alpha = 1$ , the firm's rent is no longer costly and the regulator behaves as a pure efficiency maximizer implementing the first-best output in all states of nature.

The regulation literature of the last thirty years has greatly improved our understanding of government intervention under asymmetric information. We refer to the book of Laffont and Tirole (1993) for a comprehensive view of this theory and its various implications for the design of real world regulatory institutions.

### 13.15.2 Nonlinear Pricing by a Monopoly

In Maskin and Riley (Rand J. of Economics, 1984), the principal is the seller of a private good with production cost  $cq$  who faces a continuum of buyers. The principal has thus a utility function  $V = t - cq$ . The tastes of a buyer for the private good are such that his utility function is  $U = \theta u(q) - t$ , where  $q$  is the quantity consumed and  $t$  his payment to the principal. Suppose that the parameter  $\theta$  of each buyer is drawn independently from the same distribution on  $\Theta = \{\underline{\theta}, \bar{\theta}\}$  with respective probabilities  $1 - \nu$  and  $\nu$ .

We are now in a setting with a continuum of agents. However, it is mathematically equivalent to the framework with a single agent. Now  $\nu$  is the frequency of type  $\underline{\theta}$  by the Law of Large Numbers.

Incentive and participation constraints can as usual be written directly in terms of the

information rents  $\underline{U} = \underline{\theta}u(\underline{q}) - \underline{t}$  and  $\bar{U} = \bar{\theta}u(\bar{q}) - \bar{t}$  as

$$\underline{U} \geq \bar{U} - \Delta\theta u(\bar{q}), \quad (13.84)$$

$$\bar{U} \geq \underline{U} + \Delta\theta u(\underline{q}), \quad (13.85)$$

$$\underline{U} \geq 0, \quad (13.86)$$

$$\bar{U} \geq 0. \quad (13.87)$$

The principal's program now takes the following form:

$$\max_{\{(\bar{U}, \bar{q}); (\underline{U}, \underline{q})\}} v(\bar{\theta}u(\bar{q}) + (1-v)(\underline{\theta}u(\underline{q}) - c\underline{q}) - (\nu\bar{U} + (1-\nu)\underline{U}))$$

subject to (13.84) to (13.87).

The analysis is the mirror image of that of the standard model discussed before, where now the efficient type is the one with the highest valuation for the good  $\bar{\theta}$ . Hence, (13.85) and (13.86) are the two binding constraints. As a result, there is no output distortion with respect to the first-best outcome for the high valuation type and  $\bar{q}^{SB} = \bar{q}^*$ , where  $\bar{\theta}u'(\bar{q}^*) = c$ . However, there exists a downward distortion of the low valuation agent's output with respect to the first-best outcome. We have  $\underline{q}^{SB} < \underline{q}^*$ , where

$$\left(\underline{\theta} - \frac{\nu}{1-\nu}\Delta\theta\right)u'(\underline{q}^{SB}) = c \quad \text{and} \quad \underline{\theta}u'(\underline{q}^*) = c. \quad (13.88)$$

So the unit price is not the same if the buyers demand  $\bar{q}^*$  or  $\underline{q}^{SB}$ , hence the expression of nonlinear prices.

### 13.15.3 Quality and Price Discrimination

Mussa and Rosen (JET, 1978) studied a very similar problem to the nonlinear pricing, where agents buy one unit of a commodity with quality  $q$  but are vertically differentiated with respect to their preferences for the good. The marginal cost (and average cost) of producing one unit of quality  $q$  is  $C(q)$  and the principal has the payoff function  $V = t - C(q)$ . The payoff function of an agent is now  $U = \theta q - t$  with  $\theta$  in  $\Theta = \{\underline{\theta}, \bar{\theta}\}$ , with respective probabilities  $1 - \nu$  and  $\nu$ .

Incentive and participation constraints can still be written directly in terms of the information rents  $\underline{U} = \underline{\theta}q - \underline{t}$  and  $\bar{U} = \bar{\theta}q - \bar{t}$  as

$$\underline{U} \geq \bar{U} - \Delta\theta q, \quad (13.89)$$

$$\bar{U} \geq \underline{U} + \Delta\theta\underline{q}, \quad (13.90)$$

$$\underline{U} \geq 0, \quad (13.91)$$

$$\bar{U} \geq 0. \quad (13.92)$$

The principal solves now:

$$\max_{\{(\underline{U}, \underline{q}); (\bar{U}, \bar{q})\}} v(\bar{\theta}\bar{q} - C(\bar{q})) + (1 - \nu)(\underline{\theta}\underline{q} - C(\underline{q})) - (\nu\bar{U} + (1 - \nu)\underline{U})$$

subject to (13.89) to (13.92).

Following procedures similar to what we have done so far, only (13.90) and (13.91) are binding constraints. Finally, we find that the high valuation agent receives the first-best quality  $\bar{q}^{SB} = \bar{q}^*$  where  $\bar{\theta} = C'(\bar{q}^*)$ . However, quality is now reduced below the first-best for the low valuation agent. We have  $\underline{q}^{SB} < \underline{q}^*$ , where

$$\underline{\theta} = C'(\underline{q}^{SB}) + \frac{\nu}{1 - \nu}\Delta\theta \quad \text{and} \quad \underline{\theta} = C'(\underline{q}^*) \quad (13.93)$$

Interestingly, the spectrum of qualities is larger under asymmetric information than under complete information. This incentive of the seller to put a low quality good on the market is a well-documented phenomenon in the industrial organization literature. Some authors have even argued that damaging its own goods may be part of the firm's optimal selling strategy when screening the consumers' willingness to pay for quality is an important issue.

### 13.15.4 Financial Contracts

Asymmetric information significantly affects the financial markets. For instance, in a paper by Freixas and Laffont (1990), the principal is a lender who provides a loan of size  $k$  to a borrower. Capital costs  $Rk$  to the lender since it could be invested elsewhere in the economy to earn the risk-free interest rate  $R$ . The lender has thus a payoff function  $V = t - Rk$ . The borrower makes a profit  $U = \theta f(k) - t$  where  $\theta f(k)$  is the production with  $k$  units of capital and  $t$  is the borrowers repayment to the lender. We assume that  $f' > 0$  and  $f'' < 0$ . The parameter  $\theta$  is a productivity shock drawn from  $\Theta = \{\underline{\theta}, \bar{\theta}\}$  with respective probabilities  $1 - \nu$  and  $\nu$ .



Incentive and participation constraints can again be written directly in terms of the borrower's information rents  $\underline{U} = \underline{\theta}f(\underline{k}) - \underline{t}$  and  $\bar{U} = \bar{\theta}f(\bar{k}) - \bar{t}$  as

$$\underline{U} \geq \bar{U} - \Delta\theta f(\bar{k}), \quad (13.94)$$

$$\bar{U} \geq \underline{U} + \Delta\theta f(\underline{k}), \quad (13.95)$$

$$\underline{U} \geq 0, \quad (13.96)$$

$$\bar{U} \geq 0. \quad (13.97)$$

The principal's program takes now the following form:

$$\max_{\{(\underline{U}, \underline{k}); (\bar{U}, \bar{k})\}} v(\bar{\theta}f(\bar{k}) - R\bar{k}) + (1 - \nu)(\underline{\theta}f(\underline{k}) - R\underline{k}) - (\nu\bar{U} + (1 - \nu)\underline{U})$$

subject to (13.94) to (13.97).

One can check that (13.95) and (13.96) are now the two binding constraints. As a result, there is no capital distortion with respect to the first-best outcome for the high productivity type and  $\bar{k}^{SB} = k^*$  where  $\bar{\theta}f'(\bar{k}^*) = R$ . In this case, the return on capital is equal to the risk-free interest rate. However, there also exists a downward distortion in the size of the loan given to a low productivity borrower with respect to the first-best outcome. We have  $\underline{k}^{SB} < \underline{k}^*$  where

$$\left( \underline{\theta} - \frac{\nu}{1 - \nu} \Delta\theta \right) f'(\underline{k}^{SB}) = R \quad \text{and} \quad \underline{\theta}f'(\underline{k}^*) = R. \quad (13.98)$$

### 13.15.5 Labor Contracts

Asymmetric information also undermines the relationship between a worker and the firm for which he works. In Green and Kahn (QJE, 1983) and Hart (RES, 1983), the principal is a union (or a set of workers) providing its labor force  $l$  to a firm.

The firm makes a profit  $\theta f(l) - t$ , where  $f(l)$  is the return on labor and  $t$  is the worker's payment. We assume that  $f' > 0$  and  $f'' < 0$ . The parameter  $\theta$  is a productivity shock drawn from  $\Theta = \{\underline{\theta}, \bar{\theta}\}$  with respective probabilities  $1 - \nu$  and  $\nu$ . The firm's objective is to maximize its profit  $U = \theta f(l) - t$ . Workers have a utility function defined on consumption and labor. If their disutility of labor is counted in monetary terms and all revenues from the firm are consumed, they get  $V = v(t - l)$  where  $l$  is their disutility of providing  $l$  units of labor and  $v(\cdot)$  is increasing and concave ( $v' > 0$ ,  $v'' < 0$ ).

In this context, the firm's boundaries are determined before the realization of the shock and contracting takes place *ex ante*. It should be clear that the model is similar to the one with a risk-averse principal and a risk-neutral agent. So, we know that the risk-averse union will propose a contract to the risk-neutral firm which provides full insurance and implements the first-best levels of employments  $\bar{l}$  and  $\underline{l}^*$  defined respectively by  $\bar{\theta}f'(\bar{l}^*) = 1$  and  $\underline{\theta}f'(\underline{l}^*) = 1$ .

When workers have a utility function exhibiting an income effect, the analysis will become much harder even in two-type models. For details, see Laffont and Martimort (2002).

## 13.16 The Optimal Contract with a Continuum of Types

In this section, we give a brief account of the continuum type case. Most of the principal-agent literature is written within this framework.

Reconsider the standard model with  $\theta$  in  $\Theta = [\underline{\theta}, \bar{\theta}]$ , with a cumulative distribution function  $F(\theta)$  and a density function  $f(\theta) > 0$  on  $[\underline{\theta}, \bar{\theta}]$ . Since the revelation principle is still valid with a continuum of types, and we can restrict our analysis to direct revelation mechanisms  $\{(q(\tilde{\theta}), t(\tilde{\theta}))\}$ , which are truthful, i.e., such that

$$t(\theta) - \theta q(\theta) \geq t(\tilde{\theta}) - \theta q(\tilde{\theta}) \quad \text{for any } (\theta, \tilde{\theta}) \in \Theta^2. \quad (13.99)$$

In particular, (13.99) implies

$$t(\theta) - \theta q(\theta) \geq t(\theta') - \theta q(\theta'), \quad (13.100)$$

$$t(\theta') - \theta' q(\theta') \geq t(\theta) - \theta' q(\theta) \quad \text{for all pairs } (\theta, \theta') \in \Theta^2. \quad (13.101)$$

Adding (13.100) and (13.101) we obtain

$$(\theta - \theta')(q(\theta') - q(\theta)) \geq 0. \quad (13.102)$$

Thus, incentive compatibility alone requires that the schedule of output  $q(\cdot)$  has to be nonincreasing. This implies that  $q(\cdot)$  is differentiable almost everywhere. So we will restrict the analysis to differentiable functions.

(13.99) implies that the following first-order condition for the optimal response  $\tilde{\theta}$  chosen by type  $\theta$  is satisfied

$$\dot{t}(\tilde{\theta}) - \theta \dot{q}(\tilde{\theta}) = 0. \quad (13.103)$$

For the truth to be an optimal response for all  $\theta$ , it must be the case that

$$\dot{t}(\theta) - \theta \dot{q}(\theta) = 0, \quad (13.104)$$

and (13.104) must hold for all  $\theta$  in  $\Theta$  since  $\theta$  is unknown to the principal.

It is also necessary to satisfy the local second-order condition,

$$\ddot{t}(\tilde{\theta})|_{\tilde{\theta}=\theta} - \theta \ddot{q}(\tilde{\theta})|_{\tilde{\theta}=\theta} \leq 0 \quad (13.105)$$

or

$$\ddot{t}(\theta) - \theta \ddot{q}(\theta) \leq 0. \quad (13.106)$$

But differentiating (13.104), (13.106) can be written more simply as

$$-\dot{q}(\theta) \geq 0. \quad (13.107)$$

(13.104) and (13.107) constitute the local incentive constraints, which ensure that the agent does not want to lie locally. Now we need to check that he does not want to lie globally either, therefore the following constraints must be satisfied

$$t(\theta) - \theta q(\theta) \geq t(\tilde{\theta}) - \theta q(\tilde{\theta}) \quad \text{for any } (\theta, \tilde{\theta}) \in \Theta^2. \quad (13.108)$$

From (13.104) we have

$$t(\theta) - t(\tilde{\theta}) = \int_{\tilde{\theta}}^{\theta} \tau \dot{q}(\tau) d\tau = \theta q(\theta) - \tilde{\theta} q(\tilde{\theta}) - \int_{\tilde{\theta}}^{\theta} q(\tau) d\tau \quad (13.109)$$

or

$$t(\theta) - \theta q(\theta) = t(\tilde{\theta}) - \theta q(\tilde{\theta}) + (\theta - \tilde{\theta}) q(\tilde{\theta}) - \int_{\tilde{\theta}}^{\theta} q(\tau) d\tau, \quad (13.110)$$

where  $(\theta - \tilde{\theta}) q(\tilde{\theta}) - \int_{\tilde{\theta}}^{\theta} q(\tau) d\tau \geq 0$ , because  $q(\cdot)$  is nonincreasing.

So, it turns out that the local incentive constraints (13.104) also imply the global incentive constraints.

In such circumstances, the infinity of incentive constraints (13.108) reduces to a differential equation and to a monotonicity constraint. Local analysis of incentives is enough.

Truthful revelation mechanisms are then characterized by the two conditions (13.104) and (13.107).

Let us use the rent variable  $U(\theta) = t(\theta) - \theta q(\theta)$ . The local incentive constraint is now written as (by using (13.104))

$$\dot{U}(\theta) = -q(\theta). \quad (13.111)$$

The optimization program of the principal becomes

$$\max_{\{(U(\cdot), q(\cdot))\}} \int_{\underline{\theta}}^{\bar{\theta}} (S(q(\theta)) - \theta q(\theta) - U(\theta)) f(\theta) d\theta \quad (13.112)$$

subject to

$$\dot{U}(\theta) = -q(\theta), \quad (13.113)$$

$$\dot{q}(\theta) \leq 0, \quad (13.114)$$

$$U(\theta) \geq 0. \quad (13.115)$$

Using (13.113), the participation constraint (13.115) simplifies to  $U(\bar{\theta}) \geq 0$ . As in the discrete case, incentive compatibility implies that only the participation constraint of the most inefficient type can be binding. Furthermore, it is clear from the above program that it will be binding. i.e.,  $U(\bar{\theta}) = 0$ .

Momentarily ignoring (13.114), we can solve (13.113)

$$U(\bar{\theta}) - U(\theta) = - \int_{\theta}^{\bar{\theta}} q(\tau) d\tau \quad (13.116)$$

or, since  $U(\bar{\theta}) = 0$ ,

$$U(\theta) = \int_{\theta}^{\bar{\theta}} q(\tau) d\tau \quad (13.117)$$

The principal's objective function becomes

$$\int_{\underline{\theta}}^{\bar{\theta}} \left( S(q(\theta)) - \theta q(\theta) - \int_{\theta}^{\bar{\theta}} q(\tau) d\tau \right) f(\theta) d\theta, \quad (13.118)$$

which, by an integration of parts, gives

$$\int_{\underline{\theta}}^{\bar{\theta}} \left( S(q(\theta)) - \left( \theta + \frac{F(\theta)}{f(\theta)} \right) q(\theta) \right) f(\theta) d\theta. \quad (13.119)$$

Maximizing pointwise (13.119), we get the second-best optimal outputs

$$S'(q^{SB}(\theta)) = \theta + \frac{F(\theta)}{f(\theta)}, \quad (13.120)$$

which is the first order condition for the case of a continuum of types.

If the monotone hazard rate property  $\frac{d}{d\theta} \left( \frac{F(\theta)}{f(\theta)} \right) \geq 0$  holds, the solution  $q^{SB}(\theta)$  of (13.120) is clearly decreasing, and the neglected constraint (13.114) is satisfied. All types choose therefore different allocations and there is no bunching in the optimal contract.

From (13.120), we note that there is no distortion for the most efficient type (since  $F(\underline{\theta}) = 0$  and a downward distortion for all the other types.

All types, except the least efficient one, obtain a positive information rent at the optimal contract

$$U^{SB}(\theta) = \int_{\underline{\theta}}^{\bar{\theta}} q^{SB}(\tau) d\tau. \quad (13.121)$$

Finally, one could also allow for some shutdown of types. The virtual surplus  $S(q) - \left( \theta + \frac{F(\theta)}{f(\theta)} \right) q$  decreases with  $\theta$  when the monotone hazard rate property holds, and shutdown (if any) occurs on an interval  $[\theta^*, \bar{\theta}]$ .  $\theta^*$  is obtained as a solution to

$$\max_{\{\theta^*\}} \int_{\underline{\theta}}^{\theta^*} \left( S(q^{SB}(\theta)) - \left( \theta + \frac{F(\theta)}{f(\theta)} \right) q^{SB}(\theta) \right) f(\theta) d\theta.$$

For an interior optimum, we find that

$$S(q^{SB}(\theta^*)) = \left( \theta^* + \frac{F(\theta^*)}{f(\theta^*)} \right) q^{SB}(\theta^*).$$

As in the discrete case, one can check that the Inada condition  $S'(0) = +\infty$  and the condition  $\lim_{q \rightarrow 0} S'(q)q = 0$  ensure the corner solution  $\theta^* = \bar{\theta}$ .

**Remark 13.16.1** The optimal solution above can also be derived by using the Pontryagin principle. The Hamiltonian is then

$$H(q, U, \mu, \theta) = (S(q) - \theta q - U) f(\theta) - \mu q, \quad (13.122)$$

where  $\mu$  is the co-state variable,  $U$  the state variable and  $q$  the control variable,

From the Pontryagin principle,

$$\dot{\mu}(\theta) = -\frac{\partial H}{\partial U} = f(\theta). \quad (13.123)$$

From the transversality condition (since there is no constraint on  $U(\cdot)$  at  $\underline{\theta}$ ),

$$\mu(\underline{\theta}) = 0. \quad (13.124)$$

Integrating (13.123) using (13.124), we get

$$\mu(\theta) = F(\theta). \quad (13.125)$$

Optimizing with respect to  $q(\cdot)$  also yields

$$S'(q^{SB}(\theta)) = \theta + \frac{\mu(\theta)}{f(\theta)}, \quad (13.126)$$

and inserting the value of  $\mu(\theta)$  obtained from (13.125) again yields (13.120).

We have derived the optimal truthful direct revelation mechanism  $\{(q^{SB}(\theta), U^{SB}(\theta))\}$  or  $\{(q^{SB}(\theta), t^{SB}(\theta))\}$ . It remains to be investigated if there is a simple implementation of this mechanism. Since  $q^{SB}(\cdot)$  is decreasing, we can invert this function and obtain  $\theta^{SB}(q)$ . Then,

$$t^{SB}(\theta) = U^{SB}(\theta) + \theta q^{SB}(\theta) \quad (13.127)$$

becomes

$$T(q) = t^{SB}(\theta^{SB}(q)) = \int_{\theta(q)}^{\bar{\theta}} q^{SB}(\tau) d\tau + \theta(q)q. \quad (13.128)$$

To the optimal truthful direct revelation mechanism we have associated a nonlinear transfer  $T(q)$ . We can check that the agent confronted with this nonlinear transfer chooses the same allocation as when he is faced with the optimal revelation mechanism. Indeed, we have  $\frac{d}{dq}(T(q) - \theta q) = T'(q) - \theta = \frac{dt^{SB}}{d\theta} \cdot \frac{d\theta^{SB}}{dq} - \theta = 0$ , since  $\frac{dt^{SB}}{d\theta} - \theta \frac{dq^{SB}}{d\theta} = 0$ .

To conclude, the economic insights obtained in the continuum case are not different from those obtained in the two-state case.

## 13.17 Further Extensions

The main theme of this chapter was to determine how the fundamental conflict between rent extraction and efficiency could be solved in a principal-agent relationship with adverse selection. In the models discussed, this conflict was relatively easy to understand because it resulted from the simple interaction of a single incentive constraint with a single participation constraint. Here we would mention some possible extensions.

One can consider a straightforward three-type extension of the standard model. One can also deal with a bidimensional adverse selection model, a two-type model with type-dependent reservation utilities, random participation constraints, the limited liability con-

straints, and the audit models. For detailed discussion about these topics and their applications, see Laffont and Martimort (2002).

## Reference

- Akerlof, G., "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 89 (1970), 488-500.
- Baron, D., and R. Myerson, "Regulating a Monopolist with Unknown Cost," *Econometrica*, 50 (1982), 745-782.
- Bolton, P. and M. Dewatripont (2005), *Contract Theory*, MIT Press.
- Freixas, X., J.J. Laffont, "Optimal banking Contracts," In Essays in Honor of Edmond Malinvaud, Vol. 2, *Macroeconomics*, ed. P. Champsaur et al. Cambridge: MIT Press, 1990.
- Green, J., and C. Kahn, "Wage-Employment Contracts," *Quarterly Journal of Economics*, 98 (1983), 173-188.
- Grossman, S., and O. Hart, "An Analysis of the Principal Agent," *Econometrica*, 51 (1983), 7 45.
- Hart, O., "Optimal Labor Contracts under Asymmetric Information: An Introduction," *Review of Economic Studies*, 50 (1983), 3-35.
- Hurwicz, L. (1972), "On Informational Decentralized Systems," in *Decision and Organization*, Radner, R. and C. B. McGuire, eds., in Honor of J. Marschak, (North-Holland), 297-336.
- Laffont, J.-J. and D. Martimort, *The Theory of Incentives: The Principal-Agent Model*, Princeton and Oxford: Princeton University Press, 2002, Chapters 1-3.
- Laffont, J.-J., and J. Tirole, *The Theory of Incentives in Procurement and Regulation*, Cambridge: MIT Press, 1993.
- Li, J. and G. Tian, "Optimal Contracts for Central Banks Revised," Working Paper, Texas A&M University, 2003.

- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 12.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic*, Oxford University Press, 1995, Chapter 13-14.
- Maskin, E., and J. Riley, "Monopoly with Incomplete Information," *Rand Journal of Economics*, 15 (1984), 171-196.
- Meng, D. and G. Tian, "Nonlinear Pricing with Network Externalities and Countervailing Incentives," Texas A&M University, 2008, website: <http://econ.tamu.edu/tian/paper.htm>.
- Meng, D. and G. Tian, "Multi-task Incentive Contract and Performance Measurement with Multidimensional Types," *Games and Economic Behavior*, 77 (2013) 377-404.
- Mussa, M., and S. Rosen, "Monopoly and Product Quality," *Journal of Economic Theory*, 18 (1978), 301-317.
- Rothschild, M., and J. Stiglitz, "Equilibrium in Competitive Insurance Markets," *Quarterly Journal of Economics*, 93 (1976), 541-562.
- Segal, I. (2007), *Lecture Notes in Contract Theory*, Stanford University.
- Spence, M, "Job Market Signaling," *Quarterly Journal of Economics*, 87 (1973), 355-374.
- Stiglitz, J., "Monopoly Non Linear Pricing and Imperfect Information: The Insurance Market," *Review of Economic Studies*, 44 (1977), 407-430.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 25.
- Williamson, O.E., *Markets and Hierarchies: Analysis and Antitrust Implications*, the Free Press: New York, 1975, .
- Wolfstetter, E., *Topics in Microeconomics - Industrial Organization, Auctions, and Incentives*, Cambridge Press, 1999, Chapters 8-10.



# Chapter 14

## Optimal Mechanism Design: Contracts with One-Agent and Hidden Action

### 14.1 Introduction

In the previous chapter, we stressed that the delegation of tasks creates an information gap between the principal and his agent when the latter learns some piece of information relevant to determining the efficient volume of trade. Adverse selection is not the only informational problem one can imagine. Agents may also choose actions that affect the value of trade or, more generally, the agent's performance. The principal often loses any ability to control those actions that are no longer observable, either by the principal who offers the contract or by the court of law that enforces it. In such cases we will say that there is moral hazard.

The leading candidates for such moral hazard actions are effort variables, which positively influence the agent's level of production but also create a disutility for the agent. For instance the yield of a field depends on the amount of time that the tenant has spent selecting the best crops, or the quality of their harvesting. Similarly, the probability that a driver has a car crash depends on how safely he drives, which also affects his demand for insurance. Also, a regulated firm may have to perform a costly and nonobservable investment to reduce its cost of producing a socially valuable good.

As in the case of adverse selection, asymmetric information also plays a crucial role in the design of the optimal incentive contract under moral hazard. However, instead of being an exogenous uncertainty for the principal, uncertainty is now endogenous. The probabilities of the different states of nature, and thus the expected volume of trade, now depend explicitly on the agent's effort. In other words, the realized production level is only a noisy signal of the agent's action. This uncertainty is key to understanding the contractual problem under moral hazard. If the mapping between effort and performance were completely deterministic, the principal and the court of law would have no difficulty in inferring the agent's effort from the observed output. Even if the agent's effort was not observable directly, it could be indirectly contracted upon, since output would itself be observable and verifiable.

We will study the properties of incentive schemes that induce a positive and costly effort. Such schemes must thus satisfy an incentive constraint and the agent's participation constraint. Among such schemes, the principal prefers the one that implements the positive level of effort at minimal cost. This cost minimization yields the characterization of the second-best cost of implementing this effort. In general, this second-best cost is greater than the first-best cost that would be obtained by assuming that effort is observable. An allocative inefficiency emerges as the result of the conflict of interests between the principal and the agent.

## 14.2 Basic Settings of Principal-Agent Model with Moral Hazard

### 14.2.1 Effort and Production

We consider an agent who can exert a costly effort  $e$ . Two possible values can be taken by  $e$ , which we normalize as a zero effort level and a positive effort of one:  $e$  in  $\{0, 1\}$ . Exerting effort  $e$  implies a disutility for the agent that is equal to  $\psi(e)$  with the normalization  $\psi(0) = \psi_0 = 0$  and  $\psi_1 = \psi$ .

The agent receives a transfer  $t$  from the principal. We assume that his utility function is separable between money and effort,  $U = u(t) - \psi(e)$ , with  $u(\cdot)$  increasing and concave

( $u' > 0, u'' < 0$ ). Sometimes we will use the function  $h = u^{-1}$ , the inverse function of  $u(\cdot)$ , which is increasing and convex ( $h' > 0, h'' > 0$ ).

Production is stochastic, and effort affects the production level as follows: the stochastic production level  $\tilde{q}$  can only take two values  $\{\underline{q}, \bar{q}\}$ , with  $\bar{q} - \underline{q} = \Delta q > 0$ , and the stochastic influence of effort on production is characterized by the probabilities  $\Pr(\tilde{q} = \bar{q}|e = 0) = \pi_0$ , and  $\Pr(\tilde{q} = \bar{q}|e = 1) = \pi_1$ , with  $\pi_1 > \pi_0$ . We will denote the difference between these two probabilities by  $\Delta\pi = \pi_1 - \pi_0$ .

Note that effort improves production in the sense of first-order stochastic dominance, i.e.,  $\Pr(\tilde{q} \leq q^*|e)$  is decreasing with  $e$  for any given production  $q^*$ . Indeed, we have  $\Pr(\tilde{q} \leq \underline{q}|e = 1) = 1 - \pi_1 < 1 - \pi_0 = \Pr(\tilde{q} \leq \underline{q}|e = 0)$  and  $\Pr(\tilde{q} \leq \bar{q}|e = 1) = 1 = \Pr(\tilde{q} \leq \bar{q}|e = 0)$ .

### 14.2.2 Incentive Feasible Contracts

Since the agent's action is not directly observable by the principal, the principal can only offer a contract based on the observable and verifiable production level, i.e., a function  $\{t(\tilde{q})\}$  linking the agent's compensation to the random output  $\tilde{q}$ . With two possible outcomes  $\bar{q}$  and  $\underline{q}$ , the contract can be defined equivalently by a pair of transfers  $\bar{t}$  and  $\underline{t}$ . Transfer  $\bar{t}$  (resp.  $\underline{t}$ ) is the payment received by the agent if the production  $\bar{q}$  (resp.  $\underline{q}$ ) is realized.

The risk-neutral principal's expected utility is now written as

$$V_1 = \pi_1(S(\bar{q}) - \bar{t}) + (1 - \pi_1)(S(\underline{q}) - \underline{t}) \quad (14.1)$$

if the agent makes a positive effort ( $e = 1$ ) and

$$V_0 = \pi_0(S(\bar{q}) - \bar{t}) + (1 - \pi_0)(S(\underline{q}) - \underline{t}) \quad (14.2)$$

if the agent makes no effort ( $e = 0$ ). For notational simplicity, we will denote the principal's benefits in each state of nature by  $S(\bar{q}) = \bar{S}$  and  $S(\underline{q}) = \underline{S}$ .

Each level of effort that the principal wishes to induce corresponds to a set of contracts ensuring moral hazard incentive compatibility constraint and participation constraint are satisfied:

$$\pi_1 u(\bar{t}) + (1 - \pi_1)u(\underline{t}) - \psi \geq \pi_0 u(\bar{t}) + (1 - \pi_0)u(\underline{t}) \quad (14.3)$$

$$\pi_1 u(\bar{t}) + (1 - \pi_1)u(\underline{t}) - \psi \geq 0. \quad (14.4)$$

Note that the participation constraint is ensured at the ex ante stage, i.e., before the realization of the production shock.

**Definition 14.2.1** *An incentive feasible contract satisfies the incentive compatibility and participation constraints (14.3) and (14.4).*

The timing of the contracting game under moral hazard is summarized in the figure below.

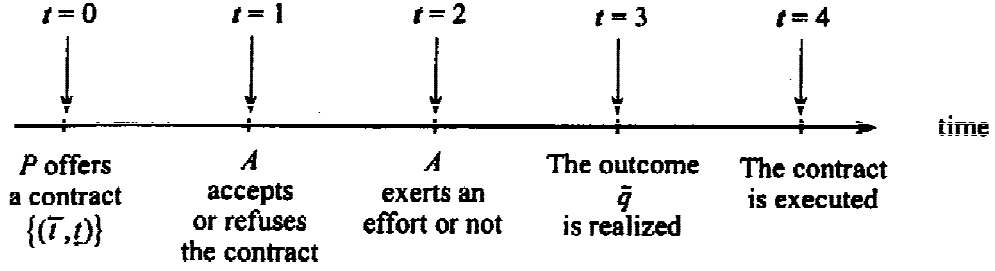


Figure 14.1: Timing of contracting under moral harzard.

### 14.2.3 The Complete Information Optimal Contract

As a benchmark, let us first assume that the principal and a benevolent court of law can both observe effort. Then, if he wants to induce effort, the principal's problem becomes

$$\max_{\{(\bar{t}, \underline{t})\}} \pi_1(\bar{S} - \bar{t}) + (1 - \pi_1)(\underline{S} - \underline{t}) \quad (14.5)$$

subject to (14.4).

Indeed, only the agents participation constraint matters for the principal, because the agent can be forced to exert a positive level of effort. If the agent were not choosing this level of effort, the agent could be heavily punished, and the court of law could commit to enforce such a punishment.

Denoting the multiplier of this participation constraint by  $\lambda$  and optimizing with respect to  $\bar{t}$  and  $\underline{t}$  yields, respectively, the following first-order conditions:

$$-\pi_1 + \lambda\pi_1 u'(\bar{t}^*) = 0, \quad (14.6)$$

$$-(1 - \pi_1) + \lambda(1 - \pi_1)u'(\underline{t}^*) = 0, \quad (14.7)$$

where  $\bar{t}^*$  and  $\underline{t}^*$  are the first-best transfers.

From (14.6) and (14.7) we immediately derive that  $\lambda = \frac{1}{u'(\underline{t}^*)} = \frac{1}{u'(\bar{t}^*)} > 0$ , and finally that  $t^* = \bar{t}^* = \underline{t}^*$ .

Thus, with a verifiable effort, the agent obtains full insurance from the risk-neutral principal, and the transfer  $t^*$  he receives is the same whatever the state of nature. Because the participation constraint is binding we also obtain the value of this transfer, which is just enough to cover the disutility of effort, namely  $t^* = h(\psi)$ . This is also the expected payment made by the principal to the agent, or the first-best cost  $C^{FB}$  of implementing the positive effort level.

For the principal, inducing effort yields an expected payoff equal to

$$V_1 = \pi_1 \bar{S} + (1 - \pi_1) \underline{S} - h(\psi) \quad (14.8)$$

Had the principal decided to let the agent exert no effort,  $e_0$ , he would make a zero payment to the agent whatever the realization of output. In this scenario, the principal would instead obtain a payoff equal to

$$V_0 = \pi_0 \bar{S} + (1 - \pi_0) \underline{S}. \quad (14.9)$$

Inducing effort is thus optimal from the principal's point of view when  $V_1 \geq V_0$ , i.e.,  $\pi_1 \bar{S} + (1 - \pi_1) \underline{S} - h(\psi) \geq \pi_0 \bar{S} + (1 - \pi_0) \underline{S}$ , or to put it differently, when the expected gain of effect is greater than first-best cost of inducing effect, i.e.,

$$\underbrace{\Delta \pi \Delta S} \geq \underbrace{h(\psi)} \quad (14.10)$$

where  $\Delta S = \bar{S} - \underline{S} > 0$ .

Denoting the benefit of inducing a strictly positive effort level by  $B = \Delta \pi \Delta S$ , the first-best outcome calls for  $e^* = 1$  if and only if  $B > h(\psi)$ , as shown in the figure below.

### 14.3 Risk Neutrality and First-Best Implementation

If the agent is risk-neutral, we have (up to an affine transformation)  $u(t) = t$  for all  $t$  and  $h(u) = u$  for all  $u$ . The principal who wants to induce effort must thus choose the contract that solves the following problem:

$$\max_{\{(\bar{t}, \underline{t})\}} \pi_1 (\bar{S} - \bar{t}) + (1 - \pi_1) (\underline{S} - \underline{t})$$

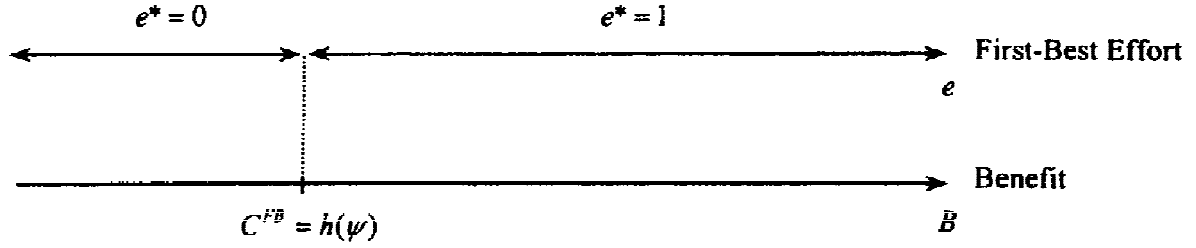


Figure 14.2: First-best level of effort.

$$\pi_1 \bar{t} + (1 - \pi_1) \underline{t} - \psi \geq \pi_0 \bar{t} + (1 - \pi_0) \underline{t} \quad (14.11)$$

$$\pi_1 \bar{t} + (1 - \pi_1) \underline{t} - \psi \geq 0. \quad (14.12)$$

With risk neutrality the principal can, for instance, choose incentive compatible transfers  $\bar{t}$  and  $\underline{t}$ , which make the agent's participation constraint binding and leave no rent to the agent. Indeed, solving (14.11) and (14.12) with equalities, we immediately obtain

$$\underline{t}^* = -\frac{\pi_0}{\Delta\pi} \psi \quad (14.13)$$

and

$$\bar{t}^* = \frac{1 - \pi_0}{\Delta\pi} \psi = \underline{t}^* + \frac{1}{\Delta\pi} \psi. \quad (14.14)$$

The agent is rewarded if production is high. His net utility in this state of nature  $\bar{U}^* = \bar{t}^* - \psi = \frac{1 - \pi_1}{\Delta\pi} \psi > 0$ . Conversely, the agent is punished if production is low. His corresponding net utility  $\underline{U}^* = \underline{t}^* - \psi = -\frac{\pi_1}{\Delta\pi} \psi < 0$ .

The principal makes an expected payment  $\pi_1 \bar{t}^* + (1 - \pi_1) \underline{t}^* = \psi$ , which is equal to the disutility of effort he would incur if he could control the effort level perfectly. The principal can costlessly structure the agent's payment so that the latter has the right incentives to exert effort. Using (14.13) and (14.14), his expected gain from exerting effort is thus  $\Delta\pi(\bar{t}^* - \underline{t}^*) = \psi$  when increasing his effort from  $e = 0$  to  $e = 1$ .

**Proposition 14.3.1** *Moral hazard is not an issue with a risk-neutral agent despite the nonobservability of effort. The first-best level of effort is still implemented.*

**Remark 14.3.1** One may find the similarity of these results with those described last chapter. In both cases, when contracting takes place ex ante, the incentive constraint,

under either adverse selection or moral hazard, does not conflict with the ex ante participation constraint with a risk-neutral agent, and the first-best outcome is still implemented.

**Remark 14.3.2** Inefficiencies in effort provision due to moral hazard will arise when the agent is no longer risk-neutral. There are two alternative ways to model these transaction costs. One is to maintain risk neutrality for positive income levels but to impose a limited liability constraint, which requires transfers not to be too negative. The other is to let the agent be strictly risk-averse. In the following, we analyze these two contractual environments and the different trade-offs they imply.

## 14.4 The Trade-Off Between Limited Liability Rent Extraction and Efficiency

Let us consider a risk-neutral agent. As we have already seen, (14.3) and (14.4) now take the following forms:

$$\pi_1 \bar{t} + (1 - \pi_1) \underline{t} - \psi \geq \pi_0 \bar{t} + (1 - \pi_0) \underline{t} \quad (14.15)$$

and

$$\pi_1 \bar{t} + (1 - \pi_1) \underline{t} - \psi \geq 0. \quad (14.16)$$

Let us also assume that the agent's transfer must always be greater than some exogenous level  $-l$ , with  $l \geq 0$ . Thus, limited liability constraints in both states of nature are written as

$$\bar{t} \geq -l \quad (14.17)$$

and

$$\underline{t} \geq -l. \quad (14.18)$$

These constraints may prevent the principal from implementing the first-best level of effort even if the agent is risk-neutral. Indeed, when he wants to induce a high effort, the principal's program is written as

$$\max_{\{(\bar{t}, \underline{t})\}} \pi_1 (\bar{S} - \bar{t}) + (1 - \pi_1) (\underline{S} - \underline{t}) \quad (14.19)$$

subject to (14.15) to (14.18).

Then, we have the following proposition.

**Proposition 14.4.1** *With limited liability, the optimal contract inducing effort from the agent entails:*

(1) *For  $l > \frac{\pi_0}{\Delta\pi}\psi$ , only (14.15) and (14.16) are binding. Optimal transfers are given by (14.13) and (14.14). The agent has no expected limited liability rent;  $EU^{SB} = 0$ .*

(2) *For  $0 \leq l \leq \frac{\pi_0}{\Delta\pi}\psi$ , (14.15) and (14.18) are binding. Optimal transfers are then given by:*

$$\underline{t}^{SB} = -l, \quad (14.20)$$

$$\bar{t}^{SB} = -l + \frac{\psi}{\Delta\pi}. \quad (14.21)$$

(3) *Moreover, the agent's expected limited liability rent  $EU^{SB}$  is non-negative:*

$$EU^{SB} = \pi_1 \bar{t}^{SB} + (1 - \pi_1) \underline{t}^{SB} - \psi = -l + \frac{\pi_0}{\Delta\pi}\psi \geq 0. \quad (14.22)$$

Proof. First suppose that  $0 \leq l \leq \frac{\pi_0}{\Delta\pi}\psi$ . We conjecture that (14.15) and (14.18) are the only relevant constraints. Of course, since the principal is willing to minimize the payments made to the agent, both constraints must be binding. Hence,  $\underline{t}^{SB} = -l$  and  $\bar{t}^{SB} = -l + \frac{\psi}{\Delta\pi}$ . We check that (14.17) is satisfied since  $-l + \frac{\psi}{\Delta\pi} > -l$ . We also check that (14.16) is satisfied since  $\pi_1 \bar{t}^{SB} + (1 - \pi_1) \underline{t}^{SB} - \psi = -l + \frac{\pi_0}{\Delta\pi}\psi \geq 0$ .

For  $l > \frac{\pi_0}{\Delta\pi}\psi$ , note that the transfers  $\underline{t}^* = -\frac{\pi_0}{\Delta\pi}\psi$ , and  $\bar{t}^* = -\psi + \frac{(1-\pi_1)}{\Delta\pi}\psi > \underline{t}^*$  are such that both limited liability constraints (14.17) and (14.18) are strictly satisfied, and (14.15) and (14.16) are both binding. In this case, it is costless to induce a positive effort by the agent, and the first-best outcome can be implemented. The proof is completed.

Note that only the limited liability constraint in the bad state of nature may be binding. When the limited liability constraint (14.18) is binding, the principal is limited in his punishments to induce effort. The risk-neutral agent does not have enough assets to cover the punishment if  $\underline{q}$  is realized in order to induce effort provision. The principal uses rewards when a good state of nature  $\bar{q}$  is realized. As a result, the agent receives a non-negative ex ante limited liability rent described by (14.22). Compared with the case without limited liability, this rent is actually the additional payment that the principal must incur because of the conjunction of moral hazard and limited liability.



As the agent becomes endowed with more assets, i.e., as  $l$  gets larger, the conflict between moral hazard and limited liability diminishes and then disappears whenever  $l$  is large enough.

## 14.5 The Trade-Off Between Insurance and Efficiency

Now suppose the agent is risk-averse. The principal's program is written as:

$$\max_{\{(\bar{t}, \underline{t})\}} \pi_1(\bar{S} - \bar{t}) + (1 - \pi_1)(\underline{S} - \underline{t}) \quad (14.23)$$

subject to (14.3) and (14.4).

Since the principal's optimization problem may not be a concave program for which the first-order Kuhn and Tucker conditions are necessary and sufficient, we make the following change of variables. Define  $\bar{u} = u(\bar{t})$  and  $\underline{u} = u(\underline{t})$ , or equivalently let  $\bar{t} = h(\bar{u})$  and  $\underline{t} = h(\underline{u})$ . These new variables are the levels of ex post utility obtained by the agent in both states of nature. The set of incentive feasible contracts can now be described by two linear constraints:

$$\pi_1 \bar{u} + (1 - \pi_1) \underline{u} - \psi \geq \pi_0 \bar{u} + (1 - \pi_0) \underline{u}, \quad (14.24)$$

$$\pi_1 \bar{u} + (1 - \pi_1) \underline{u} - \psi \geq 0, \quad (14.25)$$

which replaces (14.3) and (14.4), respectively.

Then, the principal's program can be rewritten as

$$\max_{\{(\bar{u}, \underline{u})\}} \pi_1(\bar{S} - h(\bar{u})) + (1 - \pi_1)(\underline{S} - h(\underline{u})) \quad (14.26)$$

subject to (14.24) and (14.25).

Note that the principal's objective function is now strictly concave in  $(\bar{u}, \underline{u})$  because  $h(\cdot)$  is strictly convex. The constraints are now linear and the interior of the constrained set is obviously non-empty.

### 14.5.1 Optimal Transfers

Letting  $\lambda$  and  $\mu$  be the non-negative multipliers associated respectively with the constraints (14.24) and (14.25), the first-order conditions of this program can be expressed

as

$$-\pi_1 h'(\bar{u}^{SB}) + \lambda \Delta\pi + \mu\pi_1 = -\frac{\pi_1}{u'(\bar{t}^{SB})} + \lambda \Delta\pi + \mu\pi_1 = 0, \quad (14.27)$$

$$-(1 - \pi_1)h'(\underline{u}^{SB}) - \lambda \Delta\pi + \mu(1 - \pi_1) = -\frac{(1 - \pi_1)}{u'(\underline{t}^{SB})} - \lambda \Delta\pi + \mu(1 - \pi_1) = 0. \quad (14.28)$$

where  $\bar{t}^{SB}$  and  $\underline{t}^{SB}$  are the second-best optimal transfers. Rearranging terms, we get

$$\frac{1}{u'(\bar{t}^{SB})} = \mu + \lambda \frac{\Delta\pi}{\pi_1}, \quad (14.29)$$

$$\frac{1}{u'(\underline{t}^{SB})} = \mu - \lambda \frac{\Delta\pi}{1 - \pi_1}. \quad (14.30)$$

The four variables  $(\underline{t}^{SB}, \bar{t}^{SB}, \lambda, \mu)$  are simultaneously obtained as the solutions to the system of four equations (14.24), (14.25), (14.29), and (14.30). Multiplying (14.29) by  $\pi_1$  and (14.30) by  $1 - \pi_1$ , and then adding those two modified equations we obtain

$$\mu = \frac{\pi_1}{u'(\bar{t}^{SB})} + \frac{1 - \pi_1}{u'(\underline{t}^{SB})} > 0. \quad (14.31)$$

Hence, the participation constraint (14.16) is necessarily binding. Using (14.31) and (14.29), we also obtain

$$\lambda = \frac{\pi_1(1 - \pi_1)}{\Delta\pi} \left( \frac{1}{u'(\bar{t}^{SB})} - \frac{1}{u'(\underline{t}^{SB})} \right), \quad (14.32)$$

where  $\lambda$  must also be strictly positive. Indeed, from (14.24) we have  $\bar{u}^{SB} - \underline{u}^{SB} \geq \frac{\psi}{\Delta\pi} > 0$  and thus  $\bar{t}^{SB} > \underline{t}^{SB}$ , implying that the right-hand side of (14.32) is strictly positive since  $u'' < 0$ . Using that (14.24) and (14.25) are both binding, we can immediately obtain the values of  $u(\bar{t}^{SB})$  and  $u(\underline{t}^{SB})$  by solving a system of two equations with two unknowns.

Note that the risk-averse agent does not receive full insurance anymore. Indeed, with full insurance, the incentive compatibility constraint (14.3) can no longer be satisfied. Inducing effort requires the agent to bear some risk, the following proposition provides a summary.

**Proposition 14.5.1** *When the agent is strictly risk-averse, the optimal contract that induces effort makes both the agent's participation and incentive constraints binding. This contract does not provide full insurance. Moreover, second-best transfers are given by*

$$\bar{t}^{SB} = h \left( \psi + (1 - \pi_1) \frac{\psi}{\Delta\pi} \right) = h \left( \frac{1 - \pi_0}{\Delta\pi} \psi \right) \quad (14.33)$$

and

$$\underline{t}^{SB} = h \left( \psi - \pi_1 \frac{\psi}{\Delta\pi} \right) = h \left( -\frac{\pi_0}{\Delta\pi} \psi \right). \quad (14.34)$$

### 14.5.2 The Optimal Second-Best Effort

Let us now turn to the question of the second-best optimality of inducing a high effort, from the principal's point of view. The second-best cost  $C^{SB}$  of inducing effort under moral hazard is the expected payment made to the agent  $C^{SB} = \pi_1 \bar{t}^{SB} + (1 - \pi_1) \underline{t}^{SB}$ . Using (14.33) and (14.34), this cost is rewritten as

$$\begin{aligned} C^{SB} &= \pi_1 h \left( \psi + (1 - \pi_1) \frac{\psi}{\Delta\pi} \right) + (1 - \pi_1) h \left( \psi - \frac{\pi_1 \psi}{\Delta\pi} \right) \\ &= \pi_1 h \left( \frac{1 - \pi_0}{\Delta\pi} \psi \right) + (1 - \pi_1) h \left( -\frac{\pi_0}{\Delta\pi} \psi \right). \end{aligned} \quad (14.35)$$

The benefit of inducing effort is still  $B = \Delta\pi\Delta S$ , and a positive effort  $e^* = 1$  is the optimal choice of the principal whenever

$$\begin{aligned} \Delta\pi\Delta S \geq C^{SB} &= \pi_1 h \left( \psi + (1 - \pi_1) \frac{\psi}{\Delta\pi} \right) + (1 - \pi_1) h \left( \psi - \frac{\pi_1 \psi}{\Delta\pi} \right) \\ &= \pi_1 h \left( \frac{1 - \pi_0}{\Delta\pi} \psi \right) + (1 - \pi_1) h \left( -\frac{\pi_0}{\Delta\pi} \psi \right). \end{aligned} \quad (14.36)$$

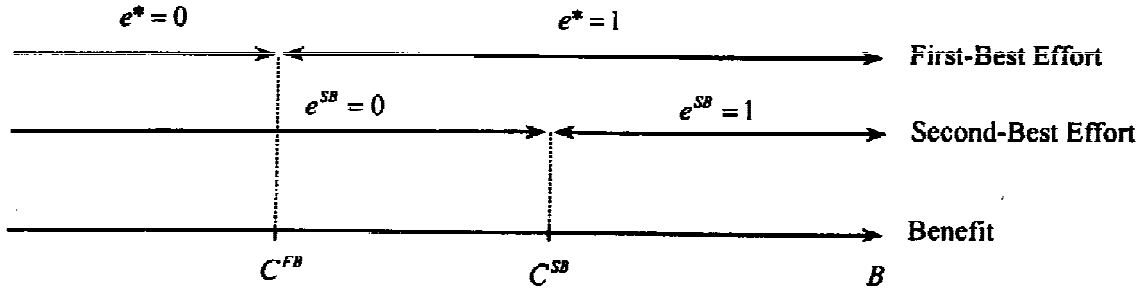


Figure 14.3: Second-best level of effort with moral hazard and risk aversion.

With  $h(\cdot)$  being strictly convex, Jensen's inequality implies that the right-hand side of (14.36) is strictly greater than the first-best cost of implementing effort  $C^{FB} = h(\psi)$ . Therefore, inducing a higher effort occurs less often with moral hazard than when effort is observable. The above figure represents this phenomenon graphically.

For  $B$  belonging to the interval  $[C^{FB}, C^{SB}]$ , the second-best level of effort is zero and is thus strictly below its first-best value. There is now an under-provision of effort because of moral hazard and risk aversion.

**Proposition 14.5.2** *With moral hazard and risk aversion, there is a trade-off between inducing effort and providing insurance to the agent. In a model with two possible levels*

of effort, the principal induces a positive effort from the agent less often than when effort is observable.

## 14.6 More than Two Levels of Performance

We now extend our previous  $2 \times 2$  model to allow for more than two levels of performance. We consider a production process where  $n$  possible outcomes can be realized. Those performances can be ordered so that  $q_1 < q_2 < \dots < q_i < \dots < q_n$ . We denote the principal's return in each of those states of nature by  $S_i = S(q_i)$ . In this context, a contract is a  $n$ -tuple of payments  $\{(t_1, \dots, t_n)\}$ . Also, let  $\pi_{ik}$  be the probability that production  $q_i$  takes place when the effort level is  $e_k$ . We assume that  $\pi_{ik}$  for all pairs  $(i, k)$  with  $\sum_{i=1}^n \pi_{ik} = 1$ . Finally, we keep the assumption that only two levels of effort are feasible. i.e.,  $e_k$  in  $\{0, 1\}$ . We still denote  $\Delta\pi_i = \pi_{i1} - \pi_{i0}$ .

### 14.6.1 Limited Liability

Consider first the limited liability model. If the optimal contract induces a positive effort, it solves the following program:

$$\max_{\{(t_1, \dots, t_n)\}} \sum_{i=1}^n \pi_{i1}(S_i - t_i) \quad (14.37)$$

subject to

$$\sum_{i=1}^n \pi_{i1}t_i - \psi \geq 0, \quad (14.38)$$

$$\sum_{i=1}^n (\pi_{i1} - \pi_{i0})t_i \geq \psi, \quad (14.39)$$

$$t_i \geq 0, \quad \text{for all } i \in \{1, \dots, n\}. \quad (14.40)$$

(14.38) is the agent's participation constraint. (14.39) is his incentive constraint. (14.40) are all the limited liability constraints by assuming that the agent cannot be given a negative payment.

First, note that the participation constraint (14.38) is implied by the incentive (14.39) and the limited liability (14.40) constraints. Indeed, we have

$$\sum_{i=1}^n \pi_{i1}t_i - \psi \geq \underbrace{\sum_{i=1}^n (\pi_{i1} - \pi_{i0})t_i - \psi}_{\geq 0} + \underbrace{\sum_{i=1}^n \pi_{i0}t_i}_{\geq 0} \geq 0.$$

Hence, we can neglect the participation constraint (14.38) in the optimization of the principal's program.

Denoting the multiplier of (14.39) by  $\lambda$  and the respective multipliers of (14.40) by  $\xi_i$ , the first-order conditions lead to

$$-\pi_{i1} + \lambda\Delta\pi_i + \xi_i = 0. \quad (14.41)$$

with the slackness conditions  $\xi_i t_i = 0$  for each  $i$  in  $\{1, \dots, n\}$ .

For such that the second-best transfer  $t_i^{SB}$  is strictly positive,  $\xi_i = 0$ , and we must have  $\lambda = \frac{\pi_{i1}}{\pi_{i1} - \pi_{i0}}$  for any such  $i$ . If the ratios  $\frac{\pi_{i1} - \pi_{i0}}{\pi_{i1}}$  all different, there exists a single index  $j$  such that  $\frac{\pi_{j1} - \pi_{j0}}{\pi_{j1}}$  is the highest possible ratio. The agent receives a strictly positive transfer only in this particular state of nature  $j$ , and this payment is such that the incentive constraint (14.39) is binding, i.e.,  $t_j^{SB} = \frac{\psi}{\pi_{j1} - \pi_{j0}}$ . In all other states, the agent receives no transfer and  $t_i^{SB} = 0$  for all  $i \neq j$ . Finally, the agent gets a strictly positive ex ante limited liability rent that is worth  $EU^{SB} = \frac{\pi_{j0}\psi}{\pi_{j1} - \pi_{j0}}$ .

The important point here is that the agent is rewarded in the state of nature that is the most informative about the fact that he has exerted a positive effort. Indeed,  $\frac{\pi_{i1} - \pi_{i0}}{\pi_{i1}}$  can be interpreted as a likelihood ratio. The principal therefore uses a maximum likelihood ratio criterion to reward the agent. The agent is only rewarded when this likelihood ratio is maximized. Like an econometrician, the principal tries to infer from the observed output what has been the parameter (effort) underlying this distribution. But here the parameter is endogenously affected by the incentive contract.

**Definition 14.6.1** *The probabilities of success satisfy the monotone likelihood ratio property (MLRP) if  $\frac{\pi_{i1} - \pi_{i0}}{\pi_{i1}}$  is nondecreasing in  $i$ .*

**Proposition 14.6.1** *If the probability of success satisfies MLRP, the second-best payment  $t_i^{SB}$  received by the agent may be chosen to be nondecreasing with the level of production  $q_i$ .*

### 14.6.2 Risk Aversion

Suppose now that the agent is strictly risk-averse. The optimal contract that induces effort must solve the program below:

$$\max_{\{t_1, \dots, t_n\}} \sum_{i=1}^n \pi_{i1}(S_i - t_i) \quad (14.42)$$

subject to

$$\sum_{i=1}^n \pi_{i1}u(t_i) - \psi \geq \sum_{i=1}^n \pi_{i0}u(t_i) \quad (14.43)$$

and

$$\sum_{i=1}^n \pi_{i1}u(t_i) - \psi \geq 0, \quad (14.44)$$

where the latter constraint is the agent's participation constraint.

Using the same change of variables as before, it should be clear that the program is again a concave problem with respect to the new variables  $u_i = u(t_i)$ . Using the same notations as before, the first-order conditions of the principal's program are written as:

$$\frac{1}{u'(t_i^{SB})} = \mu + \lambda \left( \frac{\pi_{i1} - \pi_{i0}}{\pi_{i1}} \right) \quad \text{for all } i \in \{1, \dots, n\}. \quad (14.45)$$

Multiplying each of these equations by  $\pi_{i1}$  and summing over  $i$  yields  $\mu = E_q \left( \frac{1}{u'(t_i^{SB})} \right) > 0$ , where  $E_q$  denotes the expectation operator with respect to the distribution of outputs induced by effort  $e = 1$ .

Multiplying (14.45) by  $\pi_{i1}u(t_i^{SB})$ , summing all these equations over  $i$ , and taking into account the expression of  $\mu$  obtained above yields

$$\lambda \left( \sum_{i=1}^n (\pi_{i1} - \pi_{i0})u(t_i^{SB}) \right) = E_q \left( u(\tilde{t}_i^{SB}) \left( \frac{1}{u'(\tilde{t}_i^{SB})} - E \left( \frac{1}{u'(\tilde{t}_i^{SB})} \right) \right) \right). \quad (14.46)$$

Using the slackness condition  $\lambda (\sum_{i=1}^n (\pi_{i1} - \pi_{i0})u(t_i^{SB}) - \psi) = 0$  to simplify the left-hand side of (14.46), we finally get

$$\lambda \psi = \text{cov} \left( u(\tilde{t}_i^{SB}), \frac{1}{u'(\tilde{t}_i^{SB})} \right). \quad (14.47)$$

By assumption,  $u(\cdot)$  and  $u'(\cdot)$  covary in opposite directions. Moreover, a constant wage  $t_i^{SB} = t^{SB}$  for all  $i$  does not satisfy the incentive constraint, and thus  $t_i^{SB}$  cannot be constant everywhere. Hence, the right-hand side of (14.47) is necessarily strictly positive. Thus we have  $\lambda > 0$ , and the incentive constraint is binding.

Coming back to (14.45), we observe that the left-hand side is increasing in  $t_i^{SB}$  since  $u(\cdot)$  is concave. For  $t_i^{SB}$  to be nondecreasing with  $i$ , MLRP must again hold. Then higher outputs are also those that are the more informative ones about the realization of a high effort. Hence, the agent should be more rewarded as output increases.

## 14.7 Contract Theory at Work

This section elaborates on the moral hazard paradigm discussed so far in a number of settings that have been discussed extensively in the contracting literature.

### 14.7.1 Efficiency Wage

Let us consider a risk-neutral agent working for a firm, the principal. This is a basic model studied by Shapiro and Stiglitz (AER, 1984). By exerting effort  $e$  in  $\{0, 1\}$ , the firm's added value is  $\bar{V}$  (resp.  $\underline{V}$ ) with probability  $\pi(e)$  (resp.  $1 - \pi(e)$ ). The agent can only be rewarded for a good performance and cannot be punished for a bad outcome, since they are protected by limited liability.

To induce effort, the principal must find an optimal compensation scheme  $\{(\underline{t}, \bar{t})\}$  that is the solution to the program below:

$$\max_{\{(\underline{t}, \bar{t})\}} \pi_1(\bar{V} - \bar{t}) + (1 - \pi_1)(\underline{V} - \underline{t}) \quad (14.48)$$

subject to

$$\pi_1 \bar{t} + (1 - \pi_1) \underline{t} - \psi \geq \pi_0 \bar{t} + (1 - \pi_0) \underline{t}, \quad (14.49)$$

$$\pi_1 \bar{t} + (1 - \pi_1) \underline{t} - \psi \geq 0, \quad (14.50)$$

$$\underline{t} \geq 0. \quad (14.51)$$

The problem is completely isomorphic to the one analyzed earlier. The limited liability constraint is binding at the optimum, and the firm chooses to induce a high effort when  $\Delta\pi\Delta V \geq \frac{\pi_1\psi}{\Delta\pi}$ . At the optimum,  $\underline{t}^{SB} = 0$  and  $\bar{t}^{SB} > 0$ . The positive wage  $\bar{t}^{SB} = \frac{\psi}{\Delta\pi}$  is often called an efficiency wage because it induces the agent to exert a high (efficient) level of effort. To induce production, the principal must give up a positive share of the firm's profit to the agent.

### 14.7.2 Sharecropping

The moral hazard paradigm has been one of the leading tools used by development economists to analyze agrarian economies. In the sharecropping model given in Stiglitz (RES, 1974), the principal is now a landlord and the agent is the landlord's tenant. By exerting an effort  $e$  in  $\{0, 1\}$ , the tenant increases (decreases) the probability  $\pi(e)$  (resp.  $1 - \pi(e)$ ) that a large  $\bar{q}$  (resp. small  $\underline{q}$ ) quantity of an agricultural product is produced. The price of this good is normalized to one so that the principal's stochastic return on the activity is also  $\bar{q}$  or  $\underline{q}$ , depending on the state of nature.

It is often the case that peasants in developing countries are subject to strong financial constraints. To model such a setting we assume that the agent is risk neutral and protected by limited liability. When he wants to induce effort, the principal's optimal contract must solve

$$\max_{\{\underline{t}, \bar{t}\}} \pi_1(\bar{q} - \bar{t}) + (1 - \pi_1)(\underline{q} - \underline{t}) \quad (14.52)$$

subject to

$$\pi_1\bar{t} + (1 - \pi_1)\underline{t} - \psi \geq \pi_0\bar{t} + (1 - \pi_0)\underline{t}, \quad (14.53)$$

$$\pi_1\bar{t} + (1 - \pi_1)\underline{t} - \psi \geq 0, \quad (14.54)$$

$$\underline{t} \geq 0. \quad (14.55)$$

The optimal contract therefore satisfies  $\underline{t}^{SB} = 0$  and  $\bar{t}^{SB} = \frac{\psi}{\Delta\pi}$ . This is again akin to an efficiency wage. The expected utilities obtained respectively by the principal and the agent are given by

$$EV^{SB} = \pi_1\bar{q} + (1 - \pi_1)\underline{q} - \frac{\pi_1\psi}{\Delta\pi}. \quad (14.56)$$

and

$$EU^{SB} = \frac{\pi_0\psi}{\Delta\pi}. \quad (14.57)$$

The flexible second-best contract described above has sometimes been criticized as not corresponding to the contractual arrangements observed in most agrarian economies. Contracts often take the form of simple linear schedules linking the tenant's production to his compensation. As an exercise, let us now analyze a simple linear sharing rule between the landlord and his tenant, with the landlord offering the agent a fixed share  $\alpha$  of the realized production. Such a sharing rule automatically satisfies the agent's limited



liability constraint, which can therefore be omitted in what follows. Formally, the optimal linear rule inducing effort must solve

$$\max_{\alpha}(1 - \alpha)(\pi_1\bar{q} + (1 - \pi_1)\underline{q}) \quad (14.58)$$

subject to

$$\alpha(\pi_1\bar{q} + (1 - \pi_1)\underline{q}) - \psi \geq \alpha(\pi_0\bar{q} + (1 - \pi_0)\underline{q}), \quad (14.59)$$

$$\alpha(\pi_1\bar{q} + (1 - \pi_1)\underline{q}) - \psi \geq 0 \quad (14.60)$$

Obviously, only (14.59) is binding at the optimum. One finds the optimal linear sharing rule to be

$$\alpha^{SB} = \frac{\psi}{\Delta\pi\Delta q}. \quad (14.61)$$

Note that  $\alpha^{SB} < 1$  because, for the agricultural activity to be a valuable venture in the first-best world, we must have  $\Delta\pi\Delta q > \psi$ . Hence, the return on the agricultural activity is shared between the principal and the agent, with high-powered incentives ( $\alpha$  close to one) being provided when the disutility of effort  $\psi$  is large or when the principal's gain from an increase of effort  $\Delta\pi\Delta q$  is small.

This sharing rule also yields the following expected utilities to the principal and the agent, respectively

$$EV_{\alpha} = \pi_1\bar{q} + (1 - \pi_1)\underline{q} - \left( \frac{\pi_1\bar{q} + (1 - \pi_1)\underline{q}}{\Delta q} \right) \frac{\psi}{\Delta\pi} \quad (14.62)$$

and

$$EU_{\alpha} = \left( \frac{\pi_1\bar{q} + (1 - \pi_1)\underline{q}}{\Delta q} \right) \frac{\psi}{\Delta\pi}. \quad (14.63)$$

Comparing (14.56) and (14.62) on the one hand and (14.57) and (14.63) on the other hand, we observe that the constant sharing rule benefits the agent but not the principal. A linear contract is less powerful than the optimal second-best contract. The former contract is an inefficient way to extract rent from the agent even if it still provides sufficient incentives to exert effort. Indeed, with a linear sharing rule, the agent always benefits from a positive return on his production, even in the worst state of nature. This positive return yields to the agent more than what is requested by the optimal second-best contract in the worst state of nature, namely zero. Punishing the agent for a bad performance is thus found to be rather difficult with a linear sharing rule.

A linear sharing rule allows the agent to keep some strictly positive rent  $EU_\alpha$ . If the space of available contracts is extended to allow for fixed fees  $\beta$ , the principal can nevertheless bring the agent down to the level of his outside opportunity by setting a fixed fee  $\beta^{SB}$  equal to  $\left(\frac{\pi_1 \bar{q} + (1-\pi_1) \underline{q}}{\Delta q}\right) \frac{\psi}{\Delta \pi} - \frac{\pi_0 \psi}{\Delta \pi}$ .

### 14.7.3 Wholesale Contracts

Let us now consider a manufacturer-retailer relationship studied in Laffont and Tirole (1993). The manufacturer supplies at constant marginal cost  $c$  an intermediate good to the risk-averse retailer, who sells this good on a final market. Demand on this market is high (resp. low)  $\bar{D}(p)$  (resp.  $\underline{D}(p)$ ) with probability  $\pi(e)$  where, again,  $e$  is in  $\{0, 1\}$  and  $p$  denotes the price for the final good. Effort  $e$  is exerted by the retailer, who can increase the probability that demand is high if after-sales services are efficiently performed. The wholesale contract consists of a retail price maintenance agreement specifying the prices  $\bar{p}$  and  $\underline{p}$  on the final market with a sharing of the profits, namely  $\{(\underline{t}, \underline{p}); (\bar{t}, \bar{p})\}$ . When he wants to induce effort, the optimal contract offered by the manufacturer solves the following problem:

$$\max_{\{(\underline{t}, \underline{p}); (\bar{t}, \bar{p})\}} \pi_1((\bar{p} - c)\bar{D}(\bar{p}) - \bar{t}) + (1 - \pi_1)((\underline{p} - c)\underline{D}(\underline{p}) - \underline{t}) \quad (14.64)$$

subject to (14.3) and (14.4).

The solution to this problem is obtained by appending the following expressions of the retail prices to the transfers given in (14.33) and (14.34):  $\bar{p}^* + \frac{\bar{D}(\bar{p}^*)}{\bar{D}'(\bar{p}^*)} = c$ , and  $\underline{p}^* + \frac{\underline{D}(\underline{p}^*)}{\underline{D}'(\underline{p}^*)} = c$ . Note that these prices are the same as those that would be chosen under complete information. The pricing rule is not affected by the incentive problem.

### 14.7.4 Financial Contracts

Moral hazard is an important issue in financial markets. In Holmstrom and Tirole (AER, 1994), it is assumed that a risk-averse entrepreneur wants to start a project that requires an initial investment worth an amount  $I$ . The entrepreneur has no cash of his own and must raise money from a bank or any other financial intermediary. The return on the project is random and equal to  $\bar{V}$  (resp.  $\underline{V}$ ) with probability  $\pi(e)$  (resp.  $1 - \pi(e)$ ), where the effort exerted by the entrepreneur  $e$  belongs to  $\{0, 1\}$ . We denote the spread of profits

by  $\Delta V = \bar{V} - \underline{V} > 0$ . The financial contract consists of repayments  $\{(\bar{z}, \underline{z})\}$ , depending upon whether the project is successful or not.

To induce effort from the borrower, the risk-neutral lender's program is written as

$$\max_{\{(\bar{z}, \underline{z})\}} \pi_1 \bar{z} + (1 - \pi_1) \underline{z} - I \quad (14.65)$$

subject to

$$\begin{aligned} & \pi_1 u(\bar{V} - \bar{z}) + (1 - \pi_1) u(\underline{V} - \underline{z}) - \psi \\ & \geq \pi_0 u(\bar{V} - \bar{z}) + (1 - \pi_0) u(\underline{V} - \underline{z}), \end{aligned} \quad (14.66)$$

$$\pi_1 u(\bar{V} - \bar{z}) + (1 - \pi_1) u(\underline{V} - \underline{z}) - \psi \geq 0. \quad (14.67)$$

Note that the project is a valuable venture if it provides the bank with a positive expected profit.

With the change of variables,  $\bar{t} = \bar{V} - \bar{z}$  and  $\underline{t} = \underline{V} - \underline{z}$ , the principal's program takes its usual form. This change of variables also highlights the fact that everything happens as if the lender was benefitting directly from the returns of the project, and then paying the agent only a fraction of the returns in the different states of nature.

Let us define the second-best cost of implementing a positive effort  $C^{SB}$ , and let us assume that  $\Delta\pi\Delta V \geq C^{SB}$ , so that the lender wants to induce a positive effort level even in a second-best environment. The lender's expected profit is worth

$$V_1 = \pi_1 \bar{V} + (1 - \pi_1) \underline{V} - C^{SB} - I. \quad (14.68)$$

Let us now parameterize projects according to the size of the investment  $I$ . Only the projects with positive value  $V_1 > 0$  will be financed. This requires the investment to be low enough, and typically we must have

$$I < I^{SB} = \pi_1 \bar{V} + (1 - \pi_1) \underline{V} - C^{SB}. \quad (14.69)$$

Under complete information and no moral hazard, the project would instead be financed as soon as

$$I < I^* = \pi_1 \bar{V} + (1 - \pi_1) \underline{V} \quad (14.70)$$

For intermediary values of the investment. i.e., for  $I$  in  $[I^{SB}, I^*]$ , moral hazard implies that some projects are financed under complete information but no longer under moral hazard. This is akin to some form of credit rationing.

Finally, note that the optimal financial contract offered to the risk-averse and cashless entrepreneur does not satisfy the limited liability constraint  $\underline{t} \geq 0$ . Indeed, we have  $\underline{t}^{SB} = h\left(\psi - \frac{\pi_1 \psi}{\Delta\pi}\right) < 0$ . To be induced to make an effort, the agent must bear some risk, which implies a negative payoff in the bad state of nature. Adding the limited liability constraint, the optimal contract would instead entail  $\underline{t}^{LL} = 0$  and  $\bar{t}^{LL} = h\left(\frac{\psi}{\Delta\pi}\right)$ . Interestingly, this contract has sometimes been interpreted in the corporate finance literature as a debt contract, with no money being left to the borrower in the bad state of nature and the residual being pocketed by the lender in the good state of nature.

Finally, note that

$$\begin{aligned} \bar{t}^{LL} - \underline{t}^{LL} &= h\left(\frac{\psi}{\Delta\pi}\right) < \bar{t}^{SB} - \underline{t}^{SB} = h\left(\psi + (1 - \pi_1)\frac{\psi}{\Delta\pi}\right) \\ &\quad - h\left(\psi - \frac{\pi_1 \psi}{\Delta\pi}\right), \end{aligned} \quad (14.71)$$

since  $h(\cdot)$  is strictly convex and  $h(0) = 0$ . This inequality shows that the debt contract has less incentive power than the optimal incentive contract. Indeed, it becomes harder to spread the agent's payments between both states of nature to induce effort if the agent is protected by limited liability by the agent, who is interested only in his payoff in the high state of nature, only rewards are attractive.

## 14.8 A Continuum of Performances

Let us now assume that the level of performance  $\tilde{q}$  is drawn from a continuous distribution with a cumulative function  $F(\cdot|e)$  on the support  $[\underline{q}, \bar{q}]$ . This distribution is conditional on the agent's level of effort, which still takes two possible values  $e$  in  $\{0, 1\}$ . We denote by  $f(\cdot|e)$  the density corresponding to the above distributions. A contract  $t(q)$  inducing a positive effort in this context must satisfy the incentive constraint

$$\int_{\underline{q}}^{\bar{q}} u(t(q))f(q|1)dq - \psi \geq \int_{\underline{q}}^{\bar{q}} u(t(q))f(q|0)dq, \quad (14.72)$$

and the participation constraint

$$\int_{\underline{q}}^{\bar{q}} u(t(q))f(q|1)dq - \psi \geq 0. \quad (14.73)$$

The risk-neutral principal problem is thus written as

$$\max_{\{t(q)\}} \int_{\underline{q}}^{\bar{q}} (S(q) - t(q))f(q|1)dq, \quad (14.74)$$

subject to (14.72) and (14.73).

Denoting the multipliers of (14.72) and (14.73) by  $\lambda$  and  $\mu$ , respectively, the Lagrangian is written as

$$L(q, t) = (S(q) - t)f(q|1) + \lambda(u(t)(f(q|1) - f(q|0)) - \psi) + \mu(u(t)f(q|1) - \psi).$$

Optimizing pointwise with respect to  $t$  yields

$$\frac{1}{u'(t^{SB}(q))} = \mu + \lambda \left( \frac{f(q|1) - f(q|0)}{f(q|1)} \right). \quad (14.75)$$

Multiplying (14.75) by  $f_1(q)$  and taking expectations, we obtain, as in the main text,

$$\mu = E_{\tilde{q}} \left( \frac{1}{u'(t^{SB}(\tilde{q}))} \right) > 0, \quad (14.76)$$

where  $E_{\tilde{q}}(\cdot)$  is the expectation operator with respect to the probability distribution of output induced by an effort  $e^{SB}$ . Finally, using this expression of  $\mu$ , inserting it into (14.75), and multiplying it by  $f(q|1)u(t^{SB}(q))$ , we obtain

$$\begin{aligned} & \lambda(f(q|1) - f(q|0))u(t^{SB}(q)) \\ &= f(q|1)u(t^{SB}(q)) \left( \frac{1}{u'(t^{SB}(q))} - E_{\tilde{q}} \left( \frac{1}{u'(t^{SB}(\tilde{q}))} \right) \right). \end{aligned} \quad (14.77)$$

Integrating over  $[q, \tilde{q}]$  and taking into account the slackness condition  $\lambda(\int_{\underline{q}}^{\tilde{q}} (f(q|1) - f(q|0))u(t^{SB}(q))dq - \psi) = 0$  yields  $\lambda\psi = \text{cov}(u(t^{SB}(\tilde{q})), \frac{1}{u'(t^{SB}(\tilde{q}))}) \geq 0$ .

Hence,  $\lambda \geq 0$  because  $u(\cdot)$  and  $u'(\cdot)$  vary in opposite directions. Also,  $\lambda = 0$  only if  $t^{SB}(q)$  is a constant, but in this case the incentive constraint is necessarily violated. As a result, we have  $\lambda > 0$ . Finally,  $t^{SB}(\pi)$  is monotonically increasing in  $\pi$  when the monotone likelihood property  $\frac{d}{dq} \left( \frac{f(q|1) - f(q|0)}{f(q|1)} \right) \geq 0$  is satisfied.

## 14.9 Further Extension

We have stressed the various conflicts that may appear in a moral hazard environment. The analysis of these conflicts, under both limited liability and risk aversion, was made easy because of our focus on a simple  $2 \times 2$  environment with a binary effort and two levels of performance. The simple interaction between a single incentive constraint with either a limited liability constraint or a participation constraint was quite straightforward.

When one moves away from the  $2 \times 2$  model, the analysis becomes much harder, and characterizing the optimal incentive contract is a difficult task. Examples of such complex contracting environment are abound. Effort may no longer be binary but, instead, may be better characterized as a continuous variable. A manager may no longer choose between working or not working on a project but may be able to fine-tune the exact effort spent on this project. Even worse, the agent's actions may no longer be summarized by a one-dimensional parameter but may be better described by a whole array of control variables that are technologically linked. For instance, the manager of a firm may have to choose how to allocate his effort between productive activities and monitoring his peers and other workers.

Nevertheless, one can extend the standard model to the cases where the agent can perform more than two and possibly a continuum of levels of effort, to the case with a multitask model, the case where the agent's utility function is no longer separable between consumption and effort. One can also analyze the trade-off between efficiency and redistribution in a moral hazard context. For detailed discussion, see Chapter 5 of Laffont and Martimort (2002).

## Reference

- Akerlof, G., "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 89 (1976), 400-500.
- Bolton, P. and M. Dewatripont (2005), *Contract Theory*, MIT Press.
- Grossman, S., and O. Hart, "An Analysis of the Principal Agent," *Econometrica*, 51 (1983), 7-45.
- Holmstrom, B., and J. Tirole, "Financial Intermediation, Loanable Funds, and the Real Sector," *American Economic Review*, 84 (1994), 972-991.
- Laffont, J.-J., "The New Economics of Regulation Ten Years After," *Econometrica* 62 (1994), 507-538.
- Laffont, J.-J., Laffont and D. Martimort, *The Theory of Incentives: The Principal-Agent Model*, Princeton and Oxford: Princeton University Press, 2002, Chapters 4-5.

- Laffont, J.-J., and J. Tirole, *The Theory of Incentives in Procurement and Regulation*, Cambridge: MIT Press, 1993.
- Li, J. and G. Tian, “Optimal Contracts for Central Banks Revised,” Working Paper, Texas A&M University, 2003.
- Luenberger, D. *Microeconomic Theory*, McGraw-Hill, Inc, 1995, Chapter 12.
- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic*, Oxford University Press, 1995, Chapter 14.
- Meng, D. and G. Tian, ‘Multi-task Incentive Contract and Performance Measurement with Multidimensional Types,” *Games and Economic Behavior*, 77 (2013) 377-404.
- Segal, I. (2007), *Lecture Notes in Contract Theory*, Stanford University.
- Shapiro, C., and J. Stiglitz, “Equilibrium Unemployment as a Worker Discipline Ddvice,” *American Economic Review*, 74 (1984), 433-444.
- Stiglitz, J., “Incentives and Risk Sharing in Sharecropping,” *Review of Economic Studies*, 41 (1974), 219-255.
- Varian, H.R., *Microeconomic Analysis*, W.W. Norton and Company, Third Edition, 1992, Chapters 25.
- Wolfstetter, E., *Topics in Microeconomics - Industrial Organization, Auctions, and Incentives*, Cambridge Press, 1999, Chapters 8-10.

# Chapter 15

## General Mechanism Design: Contracts with Multi-Agents

### 15.1 Introduction

In the previous chapters on the principal-agent theory, we have introduced basic models to explain the core of the principal-agent theory with complete contracts. It highlights the various trade-offs between allocative efficiency and the distribution of information rents. Since the model involves only one agent, the design of the principal's optimal contract has reduced to a constrained optimization problem without having to appeal to sophisticated game theory concepts.

In this chapter, we will introduce some of basic results and insights of the mechanism design in general, and implementation theory in particular for situations where there is one principal (also called the designer) and several agents. In such a case, asymmetric information may not only affect the relationship between the principal and each of his agents, but it may also plague the relationships between agents. To describe the strategic interaction between agents and the principal, the game theoretic reasoning is thus used to model social institutions as varied voting systems, auctions, bargaining protocols, and methods for deciding on public projects.

Incentive problems arise when the social planner cannot distinguish between things that are indeed different so that free-rider problem may appear. A free rider can improve his welfare by not telling the truth about his own un-observable characteristic. Like the



principal-agent model, a basic insight of the incentive mechanism with more than one agent is that incentive constraints should be considered coequally with resource constraints. One of the most fundamental contributions of the mechanism theory has been shown that the free-rider problem may or may not occur, depending on the kind of game (mechanism) that agents play and other game theoretical solution concepts. A theme that comes out of the literature is the difficulty of finding mechanisms compatible with individual incentives that simultaneously results in a desired social goal.

Examples of incentive mechanism design that takes strategic interactions among agents exist for a long time. An early example is the Biblical story of the famous judgement of Solomon for determining who is the real mother of a baby. Two women came before the King, disputing who was the mother of a child. The King's solution used a method of threatening to cut the lively baby in two and give half to each. One woman was willing to give up the child, but another woman agreed to cut in two. The King then made his judgement and decision: The first woman is the mother, do not kill the child and give the child to the first woman. Another example of incentive mechanism design is how to cut a pie and divide equally among all participants.

The first major development was in the work of Gibbard-Hurwicz-Satterthwaite in 1970s. When information is private, the appropriate equilibrium concept is dominant strategies. These incentives adopt the form of incentive compatibility constraints where for each agent to tell truth about their characteristics must be dominant. The fundamental conclusion of Gibbard-Hurwicz-Satterthwaite's impossibility theorem is that we have to have a trade-off between the truth-telling and Pareto efficiency (or the first best outcomes in general). Of course, if one is willing to give up Pareto efficiency, we can have a truth-telling mechanism, such as Vickery-Clark-Groves mechanism. In many cases, one can ignore the first-best or Pareto efficiency, and so one can expect the truth-telling behavior.

On the other hand, we could give up the truth-telling requirement, and want to reach Pareto efficient outcomes. When the information about the characteristics of the agents is shared by individuals but not by the designer, then the relevant equilibrium concept is the Nash equilibrium. In this situation, one can give up the truth-telling, and uses a general message space. One may design a mechanism that Nash implements Pareto efficient allocations.

We will introduce these results and such trade-offs. We will also discuss the case of incomplete information in which agents do not know each other's characteristics, and we need to consider Bayesian incentive compatible mechanism.

## 15.2 Basic Settings

Theoretical framework of the incentive mechanism design consists of five components: (1) economic environments (fundamentals of economy); (2) social choice goal to be reached; (3) economic mechanism that specifies the rules of game; (4) description of solution concept on individuals' self-interested behavior, and (5) implementation of a social choice goal (incentive-compatibility of personal interests and the social goal at equilibrium).

### 15.2.1 Economic Environments

$e_i = (Z_i, w_i, \succsim_i, Y_i)$ : economic characteristic of agent  $i$  which consists of outcome space, initial endowment if any, preference relation, and the production set if agent  $i$  is also a producer;

$e = (e_1, \dots, e_n)$ : an economy;

$E$ : The set of all priori admissible economic environments.

$U = U_1 \times \dots \times U_n$ : The set of all admissible utility functions.

$\Theta = \Theta_1 \times \Theta_2 \times \dots \times \Theta_n$ : The set of all admissible parameters  $\theta = (\theta_1, \dots, \theta_I) \in \Theta$  that determine types of parametric utility functions  $u_i(\cdot, \theta_i)$ , and so it is called the space of types or called the state of the world.

**Remark 15.2.1** Here,  $E$  is a general expression of economic environments. However, depending on the situations facing the designer, the set of admissible economic environments under consideration sometimes may be just given by  $E = U$ ,  $E = \Theta$ , or by the set of all possible initial endowments, or production sets.

The designer is assumed that he does not know individuals' economic characteristics. The individuals may or may not know the characteristics of the others. If they know, it is called the **complete information** case, otherwise it is called the **incomplete information** case.

For simplicity (but without loss of generality), when individuals' economic characteristics are private information (the case of incomplete information), we assume preferences are given by parametric utility functions. In this case, each agent  $i$  privately observes a type  $\theta_i \in \Theta_i$ , which determines his preferences over outcomes. The state  $\theta$  is drawn randomly from a prior distribution with density  $\varphi(\cdot)$  that can also be probabilities for finite  $\Theta$ . Each agent maximizes von Neumann-Morgenstern expected utility over outcomes, given by (Bernoulli) utility function  $u_i(y, \theta_i)$ . Thus, **Information structure** is specified by

- (1)  $\theta_i$  is privately observed by agent  $i$ ;
- (2)  $\{u_i(\cdot, \cdot)\}_{i=1}^n$  is common knowledge;
- (3)  $\varphi(\cdot)$  is common knowledge.

In this chapter, we will first discuss the case of complete information and then the case of incomplete information.

### 15.2.2 Social Goal

Given economic environments, each agent participates economic activities, makes decisions, receives benefits and pays costs on economic activities. The designer wants to reach some desired goal that is considered to be socially optimal by some criterion.

Let

$Z = Z_1 \times \dots \times Z_n$ : the outcome space (For example,  $Z = X \times Y$ ).

$A \subseteq Z$ : the feasible set.

$F : E \rightarrow \rightarrow A$ : the social goal or called social choice correspondence in which  $F(e)$  is the set of socially desired outcomes at the economy under some criterion of social optimality.

*Examples of Social Choice Correspondences:*

$P(e)$ : the set of Pareto efficient allocations.

$I(e)$ : the set of individual rational allocations.

$W(e)$ : the set of Walrasian allocations.

$L(e)$ : the set of Lindahl allocations.

$FA(e)$ : the set of fare allocations.

When  $F$  becomes a single-valued function, denoted by  $f$ , it is called a social choice function.

*Examples of Social Choice Functions:*

Solomon's goal.

Majority voting rule.

### 15.2.3 Economic Mechanism

Since the designer lacks the information about individuals' economic characteristics, he needs to design an appropriate incentive mechanism (contract or rules of game) to coordinate the personal interests and the social goal, i.e., under the mechanism, all individuals have incentives to choose actions which result in socially optimal outcomes when they pursue their personal interests. To do so, the designer informs how the information he collected from individuals is used to determine outcomes, that is, he first tells the rules of games. He then uses the information or actions of agents and the rules of game to determine outcomes of individuals. Thus, a mechanism consists of a message space and an outcome function. Let

$M_i$  : the message space of agent  $i$ .

$M = M_1 \times \dots \times M_n$ : the message space in which communications take place.

$m_i \in M_i$ : a message reported by agent  $i$ .

$m = (m_1, \dots, m_n) \in M$ : a profile of messages.

$h : M \rightarrow Z$ : outcome function that translates messages into outcomes.

$\Gamma = \langle M, h \rangle$ : a mechanism

That is, a mechanism consists of a message space and an outcome function.

**Remark 15.2.2** A mechanism is often also referred to as a game form. The terminology of game form distinguishes it from a game in game theory in number of ways. (1)

Mechanism design is normative analysis in contrast to game theory, which is positive economics. Game theory is important because it predicts how a given game will be played by agents. Mechanism design goes one step further: given the physical environment and the constraints faced by the designer, what goal can be realized or implemented? What mechanisms are optimal among those that are feasible? (2) The consequence of a profile of message is an outcome in mechanism design rather than a vector of utility payoffs. Of course, once the preference of the individuals are specified, then a game form or mechanism induces a conventional game. (3) The preferences of individuals in the mechanism design setting vary, while the preferences of a game takes as given. This distinction between mechanisms and games is critical. Because of this, an equilibrium (dominant strategy equilibrium) in mechanism design is much easier to exist than a game. (4) In designing mechanisms one must take into account incentive constraints in a way that personal interests are consistent to the goal that a designer want to implement it.

**Remark 15.2.3** In the implementation (incentive mechanism design) literature, one requires a mechanism be incentive compatible in the sense that personal interests are consistent with desired socially optimal outcomes even when individual agents are self-interested in their personal goals without paying much attention to the size of message. In the realization literature originated by Hurwicz (1972, 1986b), a sub-field of the mechanism literature, one also concerns the size of message space of a mechanism, and tries to find economic system to have small operation cost. The smaller a message space of a mechanism, the lower (transition) cost of operating the mechanism. For the neoclassical economies, it has been shown that competitive market economy system is the unique most efficient system that results in Pareto efficient and individually rational allocations (cf, Mount and Reiter (1974), Walker (1977), Osana (1978), Hurwicz (1986b), Jordan (1982), Tian (2004, 2005)).

#### 15.2.4 Solution Concept of Self-Interested Behavior

A basic assumption in economics is that individuals are self-interested in the sense that they pursue their personal interests. Unless they can be better off, they in general does not care about social interests. As a result, different economic environments and different rules of game will lead to different reactions of individuals, and thus each individual agent's

strategy on reaction will depend on his self-interested behavior which in turn depends on the economic environments and the mechanism.

Let  $b(e, \Gamma)$  be the set of equilibrium strategies that describes the self-interested behavior of individuals. Examples of such equilibrium solution concepts include Nash equilibrium, dominant strategy, Bayesian Nash equilibrium, etc.

Thus, given  $E$ ,  $M$ ,  $h$ , and  $b$ , the resulting equilibrium outcome is the composite function of the rules of game and the equilibrium strategy, i.e.,  $h(b(e, \Gamma))$ .

### 15.2.5 Implementation and Incentive Compatibility

In which sense can we see individuals's personal interests do not have conflicts with a social interest? We will call such problem as implementation problem. The purpose of an incentive mechanism design is to implement some desired socially optimal outcomes. Given a mechanism  $\Gamma$  and equilibrium behavior assumption  $b(e, \Gamma)$ , the implementation problem of a social choice rule  $F$  studies the relationship of the intersection state of  $F(e)$  and  $h(b(e, \Gamma))$ , which can be illustrated by the following diagram.

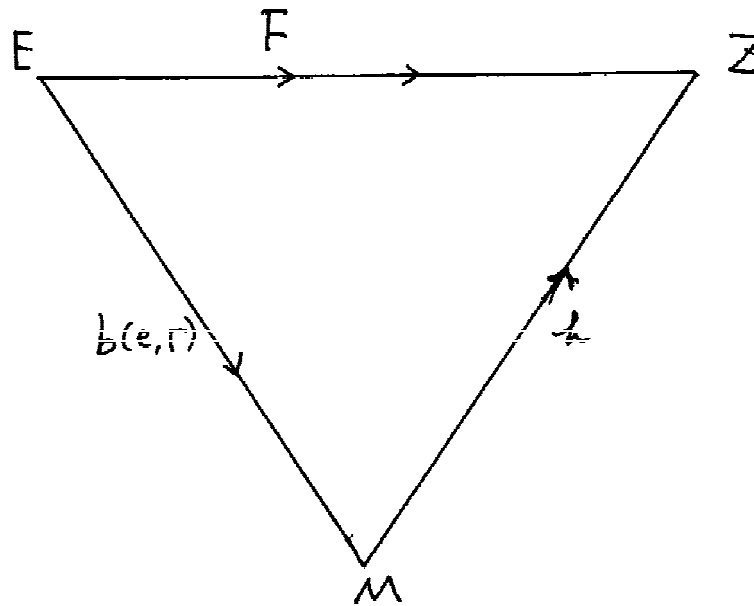


Figure 15.1: Diagrammatic Illustration of Mechanism design Problem.

We have the following various definitions on implementation and incentive compatibility of  $F$ .

A Mechanism  $\langle M, h \rangle$  is said to

- (i) fully implement a social choice correspondence  $F$  in equilibrium strategy  $b(e, \Gamma)$  on  $E$  if for every  $e \in E$

- (a)  $b(e, \Gamma) \neq \emptyset$  (equilibrium solution exists),
- (b)  $h(b(e, \Gamma)) = F(e)$  (personal interests are fully consistent with social goals);

- (ii) implement a social choice correspondence  $F$  in equilibrium strategy  $b(e, \Gamma)$  on  $E$  if for every  $e \in E$

- (a)  $b(e, \Gamma) \neq \emptyset$ ,
- (b)  $h(b(e, \Gamma)) \subseteq F(e)$ ;

- (iii) weakly implement a social choice correspondence  $F$  in equilibrium strategy  $b(e, \Gamma)$  on  $E$  if for every  $e \in E$

- (a)  $b(e, \Gamma) \neq \emptyset$ ,
- (b)  $h(b(e, \Gamma)) \cap F(e) \neq \emptyset$ .

A Mechanism  $\langle M, h \rangle$  is said to be  $b(e, \Gamma)$  *incentive-compatible with a social choice correspondence*  $F$  in  $b(e, \Gamma)$ -equilibrium if it (fully or weakly) implements  $F$  in  $b(e, \Gamma)$ -equilibrium.

Note that we did not give a specific solution concept yet when we define the implementability and incentive-compatibility. As shown in the following, whether or not a social choice correspondence is implementable will depend on the assumption on the solution concept of self-interested behavior. When information is complete, the solution concept can be dominant equilibrium, Nash equilibrium, strong Nash equilibrium, subgame perfect Nash equilibrium, undominated equilibrium, etc. For incomplete information, equilibrium strategy can be Bayesian Nash equilibrium, undominated Bayesian Nash equilibrium, etc.

## 15.3 Examples

Before we discuss some basic results in the mechanism theory, we first give some economic environments which show that one needs to design a mechanism to solve the incentive compatible problems.

**Example 15.3.1 (A Public Project)** A society is deciding on whether or not to build a public project at a cost  $c$ . The cost of the public project is shared by individuals. Let  $s_i$  be the share of the cost by  $i$  so that  $\sum_{i \in N} s_i = 1$ .

The outcome space is then  $Y = \{0, 1\}$ , where 0 represents not building the project and 1 represents building the project. Individual  $i$ 's value from use of this project is  $r_i$ . In this case, the net value of individual  $i$  is 0 from not having the project built and  $v_i = r_i - s_i c$  from having a project built. Thus agent  $i$ 's valuation function can be represented as

$$v_i(y, v_i) = y r_i - y s_i c = y v_i.$$

**Example 15.3.2 (Continuous Public Goods Setting)** In the above example, the public good could only take two values, and there is no scale problem. But, in many case, the level of public goods depends on the collection of the contribution or tax. Now let  $y \in R_+$  denote the scale of the public project and  $c(y)$  denote the cost of producing  $y$ . Thus, the outcome space is  $Z = R_+ \times R^n$ , and the feasible set is  $A = \{(y, z_1(y), \dots, z_n(y)) \in R_+ \times R^n : \sum_{i \in N} z_i(y) = c(y)\}$ , where  $z_i(y)$  is the share of agent  $i$  for producing the public goods  $y$ . The benefit of  $i$  for building  $y$  is  $r_i(y)$  with  $r_i(0) = 0$ . Thus, the net benefit of not building the project is equal to 0, the net benefit of building the project is  $r_i(y) - z_i(y)$ . The valuation function of agent  $i$  can be written as

$$v_i(y) = r_i(y) - z_i(y).$$

**Example 15.3.3 (Allocating an Indivisible Private Good)** An indivisible good is to be allocated to one member of society. For instance, the rights to an exclusive license are to be allocated or an enterprise is to be privatized. In this case, the outcome space is  $Z = \{y \in \{0, 1\}^n : \sum_{i=1}^n y_i = 1\}$ , where  $y_i = 1$  means individual  $i$  obtains the object,  $y_i = 0$  represents the individual does not get the object. If individual  $i$  gets the object, the net value benefitted from the object is  $v_i$ . If he does not get the object, his net value



is 0. Thus, agent  $i$ 's valuation function is

$$v_i(y) = v_i y_i.$$

Note that we can regard  $y$  as  $n$ -dimensional vector of public goods since  $v_i(y) = v_i y_i = v^i y$ , where  $v^i$  is a vector where the  $i$ -th component is  $v_i$  and the others are zeros, i.e.,  $v^i = (0, \dots, 0, v_i, 0, \dots, 0)$ .

From these examples, a socially optimal decision clearly depends on the individuals' true valuation function  $v_i(\cdot)$ . For instance, we have shown previously that a public project is produced if and only if the total values of all individuals is greater than its total cost, i.e., if  $\sum_{i \in N} r_i > c$ , then  $y = 1$ , and if  $\sum_{i \in N} r_i < c$ , then  $y = 0$ .

Let  $V_i$  be the set of all valuation functions  $v_i$ , let  $V = \prod_{i \in N} V_i$ , let  $h : V \rightarrow Z$  is a decision rule. Then  $h$  is said to be efficient if and only if:

$$\sum_{i \in N} v_i(h(v_i)) \geq \sum_{i \in N} v_i(h(v'_i)) \quad \forall v' \in V.$$

## 15.4 Dominant Strategy and Truthful Revelation Mechanisms

The strongest solution concept of describing self-interested behavior is dominant strategy. The dominant strategy identifies situations in which the strategy chosen by each individual is the best, regardless of choices of the others. The beauty of this equilibrium concept is in the weak rationality it demands of the agents: an agent need not forecast what the others are doing. An axiom in game theory is that agents will use it as long as a dominant strategy exists.

For  $e \in E$ , a mechanism  $\Gamma = \langle M, h \rangle$  is said to have a dominant strategy equilibrium  $m^*$  if for all  $i$

$$h_i(m_i^*, m_{-i}) \succsim_i h_i(m_i, m_{-i}) \quad \text{for all } m \in M. \quad (15.1)$$

Denote by  $D(e, \Gamma)$  the set of dominant strategy equilibria for  $\Gamma = \langle M, h \rangle$  and  $e \in E$ .

Under the assumption of dominant strategy, since each agent's optimal choice does not depend on the choices of the others and does not need to know characteristics of the others, the required information is least when an individual makes decisions. Thus, if it

exists, it is an ideal situation. Another advantage is that bad equilibria are usually not a problem. If an agent has two dominant strategies they must be payoff-equivalent, which is not a generic property.

In a given game, dominant strategies are not likely to exist. However, since we are designing the game, we can try to ensure that agents do have dominant strategies.

When the solution concept is given by dominant strategy equilibrium, i.e.,  $b(e, \Gamma) = D(e, \Gamma)$ , a mechanism  $\Gamma = \langle M, h \rangle$  implements a social choice correspondence  $F$  in dominant equilibrium strategy on  $E$  if for every  $e \in E$ ,

- (a)  $D(e, \Gamma) \neq \emptyset$ ;
- (b)  $h(D(e, \Gamma)) \subset F(e)$ .

The above definitions have applied to general (indirect) mechanisms, there is, however, a particular class of game forms which have a natural appeal and have received much attention in the literature. These are called direct or revelation mechanisms, in which the message space  $M_i$  for each agent  $i$  is the set of possible characteristics  $E_i$ . In effect, each agent reports a possible characteristic but not necessarily his true one.

A mechanism  $\Gamma = \langle M, h \rangle$  is said to be a *revelation or direct mechanism* if  $M = E$ .

**Example 15.4.1** The optimal contracts we discussed in Chapter 13 are revelation mechanisms.

**Example 15.4.2** The Groves mechanism we will discuss below is a revelation mechanism.

The most appealing revelation mechanisms are those in which truthful reporting of characteristics always turns out to be an equilibrium. It is the absence of such a mechanism which has been called the “free-rider” problem in the theory of public goods.

A revelation mechanism  $\langle E, h \rangle$  is said to implements a social choice correspondence  $F$  truthfully in  $b(e, \Gamma)$  on  $E$  if for every  $e \in E$ ,

- (a)  $e \in b(e, \Gamma)$ ;
- (b)  $h(e) \subset F(e)$ .

That is,  $F(\cdot)$  is truthfully implementable in dominant strategies if truth-telling is a dominant strategy for each agent in the direct revelation mechanism.

Truthfully implementable in dominant strategies is also called *dominant strategy incentive compatible*, *strategy proof* or *straightforward*.

Although the message space of a mechanism can be arbitrary, the following Revelation Principle tells us that one only needs to use the so-called revelation mechanism in which the message space consists solely of the set of individuals' characteristics, and it is unnecessary to seek more complicated mechanisms. Thus, it will significantly reduce the complicity of constructing a mechanism.

**Theorem 15.4.1 (Revelation Principle)** *Suppose a mechanism  $\langle M, h \rangle$  implements a social choice rule  $F$  in dominant strategy. Then there is a revelation mechanism  $\langle E, g \rangle$  which implements  $F$  truthfully in dominant strategy.*

Proof. Let  $d$  be a selection of dominant strategy correspondence of the mechanism  $\langle M, h \rangle$ , i.e., for every  $e \in E$ ,  $m^* = d(e) \in D(e, \Gamma)$ . Since  $\Gamma = \langle M, h \rangle$  implements social choice rule  $F$ , such a selection exists by the implementation of  $F$ . Since the strategy of each agent is independent of the strategies of the others, each agent  $i$ 's dominant strategy can be expressed as  $m_i^* = d_i(e_i)$ .

Define the revelation mechanism  $\langle E, g \rangle$  by  $g(e) \equiv h(d(e))$  for each  $e \in E$ . We first show that the truth-telling is always a dominant strategy equilibrium of the revelation mechanism  $\langle E, g \rangle$ . Suppose not. Then, there exists a message  $e'$  and an agent  $i$  such that

$$u_i[g(e'_i, e'_{-i})] > u_i[g(e_i, e'_{-i})].$$

However, since  $g = h \circ d$ , we have

$$u_i[h(d(e'_i), d(e'_{-i}))] > u_i[h(d(e_i), d(e'_{-i}))],$$

which contradicts the fact that  $m_i^* = d_i(e_i)$  is a dominant strategy equilibrium. This is because, when the true economic environment is  $(e_i, e'_{-i})$ , agent  $i$  has an incentive not to report  $m_i^* = d_i(e_i)$  truthfully, but have an incentive to report  $m'_i = d_i(e'_i)$ , a contradiction.

Finally, since  $m^* = d(e) \in D(e, \Gamma)$  and  $\langle M, h \rangle$  implements a social choice rule  $F$  in dominant strategy, we have  $g(e) = h(d(e)) = h(m^*) \in F(e)$ . Hence, the revelation mechanism implements  $F$  truthfully in dominant strategy. The proof is completed.

Thus, by the Revelation Principle, we know that, if truthful implementation rather than implementation is all that we require, we need never consider general mechanisms.

In the literature, if a revelation mechanism  $\langle E, h \rangle$  truthfully implements a social choice rule  $F$  in dominant strategy, the mechanism  $\Gamma$  is sometimes said to be *strongly individually incentive-compatible* with a social choice correspondence  $F$ . In particular, when  $F$  becomes a single-valued function  $f$ ,  $\langle E, f \rangle$  can be regarded as a revelation mechanism. Thus, if a mechanism  $\langle M, h \rangle$  implements  $f$  in dominant strategy, then the revelation mechanism  $\langle E, f \rangle$  is incentive compatible in dominant strategy, or called strongly individually incentive compatible.

**Remark 15.4.1** Notice that the Revelation Principle may be valid only for weak implementation. It does not apply to full implementation. The Revelation Principle specifies a correspondence between a dominant strategy equilibrium of the original mechanism  $\langle M, h \rangle$  and the true profile of characteristics as a dominant strategy equilibrium, and it does not require the revelation mechanism has a unique dominant equilibrium so that the revelation mechanism  $\langle E, g \rangle$  may also exist non-truthful strategy equilibrium that does not corresponds to any equilibrium. Thus, in moving from the general (indirect) dominant strategy mechanisms to direct ones, one may introduce undesirable dominant strategies which are not truthful. More troubling, these additional strategies may create a situation where the indirect mechanism is an implantation of a given  $F$ , while the direct revelation mechanism is not. Thus, even if a mechanism implements a social choice function, the corresponding revelation mechanism  $\langle E, g \rangle$  may only weakly implement, but not implement  $F$ .

## 15.5 Gibbard-Satterthwaite Impossibility Theorem

The Revelation Principle is very useful to find a dominant strategy mechanism. If one hopes a social choice goal  $f$  can be (weakly) implemented in dominant strategy, one only needs to show the revelation mechanism  $\langle E, f \rangle$  is strongly incentive compatible.

However, the more is known a priori about agents' characteristics, the fewer incentive-compatibility constraints an implementable choice function has to satisfy, and the more likely it is to be implementable. Thus, the worst possible case is when nothing is known about the agents' preferences over  $X$ . The Gibbard-Satterthwaite impossibility theorem in Chapter 9 tells us that, if the domain of economic environments is unrestricted, such

a mechanism does not exist unless it is a dictatorial mechanism. From the angle of the mechanism design, we restate this theorem here.

**Definition 15.5.1** A social choice function is dictatorial if there exists an agent whose optimal choice is the social optimal.

Now we state the Gibbard-Satterthwaite Theorem without the proof that is very complicated. A proof can be found, say, in Salanié's book (2000): *Microeconomics of Market Failures*.

**Theorem 15.5.1 (Gibbard-Satterthwaite Theorem)** *If  $X$  has at least 3 alternatives, a social choice function which is strongly individually incentive compatible and defined on a unrestricted domain is dictatorial.*

## 15.6 Hurwicz Impossibility Theorem

The Gibbard-Satterthwaite impossibility theorem is a very negative result. This result is essentially equivalent to Arrow's impossibility result. However, as we will show, when the admissible set of economic environments is restricted, the result may be positive as the Groves mechanism defined on quasi-linear utility functions. Unfortunately, the following Hurwicz's impossibility theorem shows the Pareto efficiency and the truthful revelation is fundamentally inconsistent even for the class of neoclassical economic environments.

**Theorem 15.6.1 (Hurwicz Impossibility Theorem, 1972)** *For the neoclassical private goods economies, there is no mechanism  $\langle M, h \rangle$  that implements Pareto efficient and individually rational allocations in dominant strategy. Consequently, any revelation mechanism  $\langle M, h \rangle$  that yields Pareto efficient and individually rational allocations is not strongly individually incentive compatible. (Truth-telling about their preferences is not Nash Equilibrium).*

Proof: By the Revelation Principle, we only need to show that any revelation mechanism cannot implement Pareto efficient and individually rational allocations truthfully in dominant equilibrium for a particular pure exchange economy. In turn, it is enough to show that truth-telling is not a Nash equilibrium for any revelation mechanism that

yields Pareto efficient and individually rational allocations for a particular pure exchange economy.

Consider a private goods economy with two agents ( $n = 2$ ) and two goods ( $L = 2$ ),

$$w_1 = (0, 2), w_2 = (2, 0)$$

$$\hat{u}_i(x, y) = \begin{cases} 3x_i + y_i & \text{if } x_i \leq y_i \\ x_i + 3y_i & \text{if } x_i > y_i \end{cases}$$

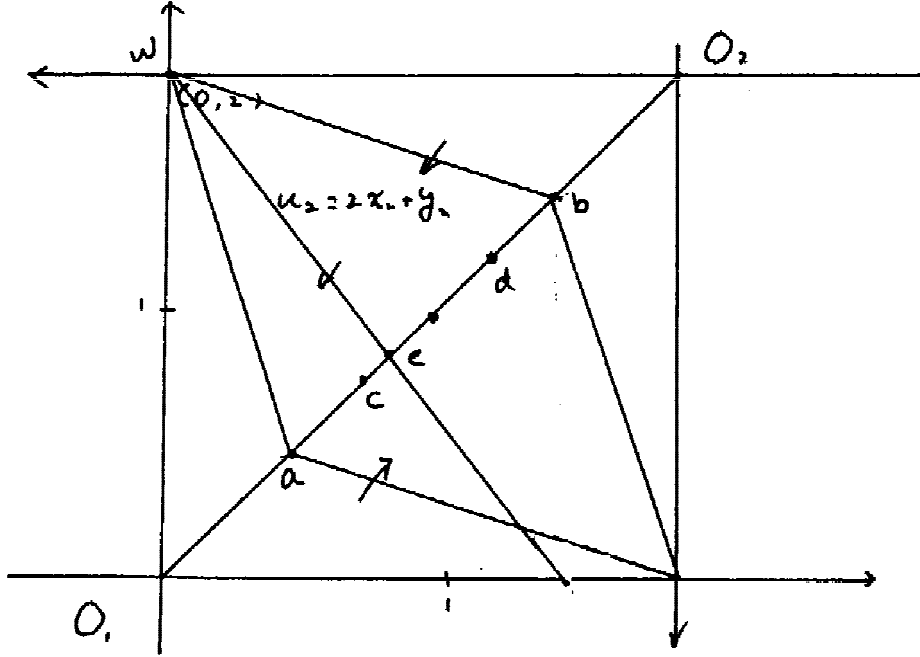


Figure 15.2: An illustration of the proof of Hurwicz's impossibility theorem.

Thus, feasible allocations are given by

$$A = \{[(x_1, y_1), (x_2, y_2)] \in R_+^4 : \\ x_1 + x_2 = 2 \\ y_1 + y_2 = 2\}$$

Let  $U_i$  be the set of all neoclassical utility functions, i.e. they are continuous and quasi-concave, which agent  $i$  can report to the designer. Thus, the true utility function  $\hat{u}_i \in U_i$ . Then,

$$U = U_1 \times U_2$$

$$h : U \rightarrow A$$

Note that, if the true utility function profile  $\hat{u}_i$  were a Nash Equilibrium, it would satisfy

$$\hat{u}_i(h_i(\hat{u}_i, \hat{u}_{-i})) \geq \hat{u}_i(h_i(u_i, \hat{u}_{-i})) \quad (15.2)$$

We want to show that  $\hat{u}_i$  is not a Nash equilibrium. Note that,

$$(1) P(e) = \overline{O_1 O_2} \text{ (contract curve)}$$

$$(2) IR(e) \cap P(e) = \overline{ab}$$

$$(3) h(\hat{u}_1, \hat{u}_2) = d \in \overline{ab}$$

Now, suppose agent 2 reports his utility function by cheating:

$$u_2(x_2, y_2) = 2x + y \quad (15.3)$$

Then, with  $u_2$ , the new set of individually rational and Pareto efficient allocations is given by

$$IR(e) \cap P(e) = \overline{ae} \quad (15.4)$$

Note that any point between  $a$  and  $e$  is strictly preferred to  $d$  by agent 2. Thus, an allocation determined by any mechanism which is IR and Pareto efficient allocation under  $(\hat{u}_1, u_2)$  is some point, say, the point  $c$  in the figure, between the segment of the line determined by  $a$  and  $e$ . Hence, we have

$$\hat{u}_2(h_2(\hat{u}_1, u_2)) > \hat{u}_2(h_2(\hat{u}_1, \hat{u}_2)) \quad (15.5)$$

since  $h_2(\hat{u}_1, u_2) = c \in \overline{ae}$ . Similarly, if  $d$  is between  $\overline{ae}$ , then agent 1 has incentive to cheat. Thus, no mechanism that yields Pareto efficient and individually rational allocations is incentive compatible. The proof is completed.

Thus, the Hurwicz's impossibility theorem implies that Pareto efficiency and the truthful revelation about individuals' characteristics are fundamentally incompatible. However, if one is willing to give up Pareto efficiency, say, one only requires the efficient provision of public goods, is it possible to find an incentive compatible mechanism which results in the Pareto efficient provision of a public good and can truthfully reveal individuals' characteristics? The answer is positive. For the class of quasi-linear utility functions, the so-called Vickrey-Clark-Groves Mechanism can be such a mechanism.

## 15.7 Vickrey-Clark-Groves Mechanisms

From Chapter 12 on public goods, we have known that public goods economies may present problems by a decentralized resource allocation mechanism because of the free-rider problem. Private provision of a public good generally results in less than an efficient amount of the public good. Voting may result in too much or too little of a public good. Are there any mechanisms that result in the “right” amount of the public good? This is a question of the incentive compatible mechanism design.

Again, we will focus on the quasi-linear environment, where all agents are known to care for money. In this environment, the results are more positive, and we can even implement the efficient decision rule (the efficient provision of public goods). For simplicity, let us first return to the model of discrete public good.

### 15.7.1 Vickrey-Clark-Groves Mechanisms for Discrete Public Good

Consider a provision problem of a discrete public good. Suppose that the economy has  $n$  agents. Let

$c$ : the cost of producing the public project.

$r_i$ : the maximum willingness to pay of  $i$ .

$s_i$ : the share of the cost by  $i$ .

$v_i = r_i - s_i c$ : the net value of  $i$ .

The public project is determined according to

$$y = \begin{cases} 1 & \text{if } \sum_{i=1}^n v_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

From the discussion in Chapter 12, it is efficient to produce the public good,  $y = 1$ , if and only if

$$\sum_{i=1}^n v_i = \sum_{i=1}^n (r_i - g_i) \geq 0.$$

Since the maximum willingness to pay for each agent,  $r_i$ , is private information and so is the net value  $v_i$ , what mechanism one should use to determine if the project is



built? One mechanism that we might use is simply to ask each agent to report his or her net value and provide the public good if and only if the sum of the reported value is positive. The problem with such a scheme is that it does not provide right incentives for individual agents to reveal their true willingness-to-pay. Individuals may have incentives to underreport their willingness-to-pay.

Thus, a question is how we can induce each agent to truthfully reveal his true value for the public good. The so-called Vickrey-Clark-Groves (VCG) mechanism gives such a mechanism.

Suppose the utility functions are quasi-linear in net increment in private good,  $x_i - w_i$ , which have the form:

$$\begin{aligned}\bar{u}_i(x_i - w_i, y) &= x_i - w_i + r_i y \\ \text{s.t. } x_i + g_i y &= w_i + t_i\end{aligned}$$

where  $t_i$  is the transfer to agent  $i$ . Then, we have

$$\begin{aligned}u_i(t_i, y) &= t_i + r_i y - g_i y \\ &= t_i + (r_i - g_i)y \\ &= t_i + v_i y.\end{aligned}$$

- Groves Mechanism:

In a Groves mechanism, agents are required to report their net values. Thus the message space of each agent  $i$  is  $M_i = \mathfrak{R}$ . The Groves mechanism is defined as follows:

$\Gamma = (M_1, \dots, M_n, t_1(\cdot), t_2(\cdot), \dots, t_n(\cdot), y(\cdot)) \equiv (M, t(\cdot), y(\cdot))$ , where

- (1)  $b_i \in M_i = R$ : each agent  $i$  reports a “bid” for the public good, i.e., report the net value of agent  $i$  which may or may not be his true net value  $v_i$ .
- (2) The level of the public good is determined by

$$y(b) = \begin{cases} 1 & \text{if } \sum_{i=1}^n b_i \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- (3) Each agent  $i$  receives a side payment (transfer)

$$t_i = \begin{cases} \sum_{j \neq i} b_j & \text{if } \sum_{i=1}^n b_i \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (15.6)$$

Then, the payoff of agent  $i$  is given by

$$\phi_i(b) = \begin{cases} v_i + t_i = v_i + \sum_{j \neq i} b_j & \text{if } \sum_{i=1}^n b_i \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (15.7)$$

We want to show that it is optimal for each agent to report the true net value,  $b_i = v_i$ , regardless of what the other agents report. That is, truth-telling is a dominant strategy equilibrium.

**Proposition 15.7.1** *The truth-telling is a dominant strategy under the Groves-Clark mechanism.*

Proof: There are two cases to be considered.

Case 1:  $v_i + \sum_{j \neq i} b_j > 0$ . Then agent  $i$  can ensure the public good is provided by reporting  $b_i = v_i$ . Indeed, if  $b_i = v_i$ , then  $\sum_{j \neq i} b_j + v_i = \sum_{i=1}^n b_i > 0$  and thus  $y = 1$ . In this case,  $\phi_i(v_i, b_{-i}) = v_i + \sum_{j \neq i} b_j > 0$ .

Case 2:  $v_i + \sum_{j \neq i} b_j \leq 0$ . Agent  $i$  can ensure that the public good is not provided by reporting  $b_i = v_i$  so that  $\sum_{i=1}^n b_i \leq 0$ . In this case,  $\phi_i(v_i, b_{-i}) = 0 \geq v_i + \sum_{j \neq i} b_j$ .

Thus, for either cases, agent  $i$  has incentives to tell the true value of  $v_i$ . Hence, it is optimal for agent  $i$  to tell the truth. There is no incentive for agent  $i$  to misrepresent his true net value regardless of what other agents do.

The above preference revelation mechanism has a major fault: the total side-payment may be very large. Thus, it is very costly to induce the agents to tell the truth.

Ideally, we would like to have a mechanism where the sum of the side-payment is equal to zero so that the feasibility condition holds, and consequently it results in Pareto efficient allocations, but in general it is impossible by Hurwicz's impossibility theorem. However, we could modify the above mechanism by asking each agent to pay a "tax", but not receive payment. Because of this "waster" tax, the allocation of public goods is still not Pareto efficient.

The basic idea of paying a tax is to add an extra amount to agent  $i$ 's side-payment,  $d_i(b_{-i})$  that depends only on what the other agents do.

**A General Groves Mechanism:** Ask each agent to pay additional tax,  $d_i(b_{-i})$ .

In this case, the transfer is given by

$$t_i(b) = \begin{cases} \sum_{j \neq i} b_j - d_i(b_{-i}) & \text{if } \sum_{i=1}^n b_i \geq 0 \\ -d_i(b_{-i}) & \text{if } \sum_{i=1}^n b_i < 0 \end{cases}$$

The payoff to agent  $i$  now takes the form:

$$\phi_i(b) = \begin{cases} v_i + t_i - d_i(b_{-i}) = v_i + \sum_{j \neq i} b_j - d_i(b_{-i}) & \text{if } \sum_{i=1}^n b_i \geq 0 \\ -d_i(b_{-i}) & \text{otherwise} \end{cases} \quad (15.8)$$

For exactly the same reason as for the mechanism above, one can prove that it is optimal for each agent  $i$  to report his true net value.

If the function  $d_i(b_{-i})$  is suitably chosen, the size of the side-payment can be significantly reduced. A nice choice was given by Clarke (1971). He suggested a particular Groves mechanism known as the *Clarke mechanism* (also called *Pivot mechanism*):

The *Pivotal Mechanism* is a special case of the general Groves Mechanism in which  $d_i(b_{-i})$  is given by

$$d_i(b_{-i}) = \begin{cases} \sum_{j \neq i} b_j & \text{if } \sum_{j \neq i} b_j \geq 0 \\ 0 & \text{if } \sum_{j \neq i} b_j < 0 \end{cases}$$

In this case, it gives

$$t_i(b) = \begin{cases} 0 & \text{if } \sum_{i=1}^n b_i \geq 0 \text{ and } \sum_{j \neq i} b_j \geq 0, \\ \sum_{j \neq i} b_j & \text{if } \sum_{i=1}^n b_i \geq 0 \text{ and } \sum_{j \neq i} b_j < 0 \\ -\sum_{j \neq i} b_j & \text{if } \sum_{i=1}^n b_j < 0 \text{ and } \sum_{j \neq i} b_j \geq 0 \\ 0 & \text{if } \sum_{i=1}^n b_j < 0 \text{ and } \sum_{j \neq i} b_j < 0 \end{cases} \quad (15.9)$$

i.e.,

$$t_i(b) = \begin{cases} -|\sum_{j \neq i} b_j| & \text{if } (\sum_{i=1}^n b_i)(\sum_{j \neq i} b_i) < 0 \\ -|\sum_{j \neq i} b_j| & \text{if } \sum_{i=1}^n b_i = 0 \text{ and } \sum_{j \neq i} b_i < 0 \\ 0 & \text{otherwise} \end{cases} \quad (15.10)$$

Therefore, the payoff of agent  $i$

$$\phi_i(b) = \begin{cases} v_i & \text{if } \sum_{i=1}^n b_i \geq 0 \text{ and } \sum_{j \neq i} b_j \geq 0 \\ v_i + \sum_{j \neq i} b_j & \text{if } \sum_{i=1}^n b_i \geq 0 \text{ and } \sum_{j \neq i} b_j < 0 \\ -\sum_{j \neq i} b_j & \text{if } \sum_{i=1}^n b_i < 0 \text{ and } \sum_{j \neq i} b_j \geq 0 \\ 0 & \text{if } \sum_{i=1}^n b_i < 0 \text{ and } \sum_{j \neq i} b_j < 0 \end{cases} \quad (15.11)$$

**Remark 15.7.1** Thus, from the transfer given in (15.10), adding in the side-payment has the effect of taxing agent  $i$  only if he changes the social decision. Such an agent is called the pivotal person. The amount of the tax agent  $i$  must pay is the amount by which agent  $i$ 's bid damages the other agents. The price that agent  $i$  must pay to change the amount of public good is equal to the harm that he imposes on the other agents.

## 15.7.2 The Vickrey-Clark-Groves Mechanisms with Continuous Public Goods

Now we are concerned with the provision of continuous public goods. Consider a public goods economy with  $n$  agents, one private good, and  $K$  public goods. Denote

$x_i$ : the consumption of the private good by  $i$ ;

$y$ : the consumption of the public goods by all individuals;

$t_i$ : transfer payment to  $i$ ;

$g_i(y)$ : the contribution made by  $i$ ;

$c(y)$ : the cost function of producing public goods  $y$  that satisfies the condition:

$$\sum g_i(y) = c(y).$$

Then, agent  $i$ 's budget constraint should satisfy

$$x_i + g_i(y) = w_i + t_i \quad (15.12)$$

and his utility functions are given by

$$\bar{u}_i(x_i - w_i, y) = x_i - w_i + u_i(y) \quad (15.13)$$

By substitution,

$$\begin{aligned} u_i(t_i, y) &= t_i + (u_i(y) - g_i(y)) \\ &\equiv t_i + v_i(y) \end{aligned}$$

where  $v_i(y)$  is called the valuation function of agent  $i$ . From the budget constraint,

$$\sum_{i=1}^n \{x_i + g_i(y)\} = \sum_{i=1}^n w_i + \sum_{i=1}^n t_i \quad (15.14)$$

we have

$$\sum_{i=1}^n x_i + c(y) = \sum_{i=1}^n w_i + \sum_{i=1}^n t_i \quad (15.15)$$

The feasibility (or balanced) condition then becomes

$$\sum_{i=1}^n t_i = 0 \quad (15.16)$$

Recall that Pareto efficient allocations are completely characterized by

$$\begin{aligned} & \max \sum a_i \bar{u}_i(x_i, y) \\ \text{s.t.} \quad & \sum_{i=1}^n x_i + c(y) = \sum_{i=1}^n w_i \end{aligned}$$

For quasi-linear utility functions it is easily seen that the weights  $a_i$  must be the same for all agents since  $a_i = \lambda$  for interior Pareto efficient allocations (no income effect), the Lagrangian multiplier, and  $y$  is thus uniquely determined for the special case of quasi-linear utility functions  $u_i(t_i, y) = t_i + v_i(y)$ . Then, the above characterization problem becomes

$$\max_{t_i, y} \left[ \sum_{i=1}^n (t_i + v_i(y)) \right] \quad (15.17)$$

or equivalently

(1)

$$\max_y \sum_{i=1}^n v_i(y)$$

(2) (feasibility condition):  $\sum_{i=1}^n t_i = 0$

Then, the Lindahl-Samuelson condition is given by:

$$\sum_{i=1}^n \frac{\partial v_i(y)}{\partial y_k} = 0.$$

that is,

$$\sum_{i=1}^n \frac{\partial u_i(y)}{\partial y_k} = \frac{\partial c(y)}{\partial y_k}$$

Thus, Pareto efficient allocations for quasi-linear utility functions are completely characterized by the Lindahl-Samuelson condition  $\sum_{i=1}^n v'_i(y) = 0$  and feasibility condition  $\sum_{i=1}^n t_i = 0$ .

In a Groves mechanism, it is supposed that each agent is asked to report the valuation function  $v_i(y)$ . Denote his reported valuation function by  $b_i(y)$ .

To get the efficient level of public goods, the government may announce that it will provide a level of public goods  $y^*$  that maximizes

$$\max_y \sum_{i=1}^n b_i(y)$$

The Groves mechanism has the form:

$$\Gamma = (V, h) \quad (15.18)$$

where  $V = V_1 \times \dots \times V_n$  : is the message space that consists of the set of all possible valuation functions with element  $b_i(y) \in V_i$ ,  $h = (t_1(b), t_2(b), \dots, t_n(b), y(b))$  are outcome functions. It is determined by:

(1) Ask each agent  $i$  to report his/her valuation function  $b_i(y)$  which may or may not be the true valuation  $v_i(y)$ .

(2) Determine  $y^*$ : the level of the public goods,  $y^* = y(b)$ , is determined by

$$\max_y \sum_{i=1}^n b_i(y) \quad (15.19)$$

(3) Determine  $t$ : transfer of agent  $i$ ,  $t_i$  is determined by

$$t_i = \sum_{j \neq i} b_j(y^*) \quad (15.20)$$

The payoff of agent  $i$  is then given by

$$\phi_i(b(y^*)) = v_i(y^*) + t_i(b) = v_i(y^*) + \sum_{j \neq i} b_j(y^*) \quad (15.21)$$

The social planner's goal is to have the optimal level  $y^*$  that solves the problem:

$$\max_y \sum_{i=1}^n b_i(y).$$

In which case is the individual's interest consistent with social planner's interest? Under the rule of this mechanism, it is optimal for each agent  $i$  to truthfully report his true valuation function  $b_i(y) = v_i(y)$  since agent  $i$  wants to maximize

$$v_i(y) + \sum_{j \neq i} b_j(y).$$

By reporting  $b_i(y) = v_i(y)$ , agent  $i$  ensures that the government will choose  $y^*$  which also maximizes his payoff while the government maximizes the social welfare. That is, individual's interest is consistent with the social interest that is determined by the Lindahl-Samuelson condition. Thus, truth-telling,  $b_i(y) = v_i(y)$ , is a dominant strategy equilibrium.

In general,  $\sum_{i=1}^n t_i(b(y)) \neq 0$ , which means that a Groves mechanism in general does not result in Pareto efficient outcomes even if it satisfies the Lindahl-Samuelson condition, i.e., it is Pareto efficient to provide the public goods.

As in the discrete case, the total transfer can be very large, just as before, they can be reduced by an appropriate side-payment. The Groves mechanism can be modified to

$$t_i(b) = \sum_{j \neq i} b_j(y) - d_i(b_{-i}).$$

The general form of the Groves Mechanism is then  $\Gamma = \langle V, t, y(b) \rangle$  such that

- (1)  $\sum_{i=1}^n b_i(y(b)) \geq \sum_{i=1}^n b_i(y)$  for  $y \in Y$ ;
- (2)  $t_i(b) = \sum_{j \neq i} b_j(y) - d_i(b_{-i})$ .

A special case of the Groves mechanism is independently described by Clark and is called the Clark mechanism (also called the pivotal mechanism) in which  $d_i(b_{-i}(y))$  is given by

$$d_i(b_{-i}) = \max_y \sum_{j \neq i} b_j(y). \quad (15.22)$$

That is, the pivotal mechanism,  $\Gamma = \langle V, t, y(b) \rangle$ , is to choose  $(y^*, t_i^*)$  such that

- (1)  $\sum_{i=1}^n b_i(y^*) \geq \sum_{i=1}^n b_i(y)$  for  $y \in Y$ ;
- (2)  $t_i(b) = \sum_{j \neq i} b_j(y^*) - \max_y \sum_{j \neq i} b_j(y)$ .

It is interesting to point out that the Clark mechanism contains the well-known Vickery auction mechanism (the second-price auction mechanism) as a special case. Under the Vickery mechanism, the highest bidding person obtains the object, and he pays the second highest bidding price. To see this, let us explore this relationship in the case of a single good auction (Example 9.3.3 in the beginning of this chapter). In this case, the outcome space is

$$Z = \{y \in \{0, 1\}^n : \sum_{i=1}^n y_i = 1\}$$

where  $y_i = 1$  implies that agent  $i$  gets the object, and  $y_i = 0$  means the person does not get the object. Agent  $i$ 's valuation function is then given by

$$v_i(y) = v_i y_i.$$

Since we can regard  $y$  as a  $n$ -dimensional vector of public goods, by the Clark mechanism above, we know that

$$y = g(b) = \{y \in Z : \max_{i=1}^n b_i y_i\} = \{y \in Z : \max_{i \in N} b_i\},$$

and the truth-telling is a dominate strategy. Thus, if  $g_i(v) = 1$ , then  $t_i(v) = \sum_{j \neq i} v_j y_j^* - \max_y \sum_{j \neq i} v_j y_j = -\max_{j \neq i} v_j$ . If  $g_i(b) = 0$ , then  $t_i(v) = 0$ . This means that the object is allocated to the individual with the highest valuation and he pays an amount equal to the second highest valuation. No other payments are made. This is exactly the outcome predicted by the Vickrey mechanism.

Do there exist other mechanisms implementing the efficient decision rule  $y^*$ ? The answer is no if valuation functions  $v(y, \cdot)$  is sufficiently "rich" in a sense. To see this, consider the class of parametric valuation functions  $v_i(y, \cdot)$  with a continuum type space  $\Theta_i = [\underline{\theta}, \bar{\theta}]$ . We first state the following lemma.

**Lemma 15.7.1** *At a solution  $(y(\cdot), t(\cdot))$ , individual rationalities  $IR_\theta$  are not binding for all  $\theta > \underline{\theta}$ , and  $IR_\theta$  is binding at  $\underline{\theta}$ .*

Proof. As in the two-type case, single-crossing property (SCP),  $IR_\theta$  and  $IC_\theta$  imply  $IR_\theta > 0$ . Now, if  $IR_\theta$  were not binding at  $\underline{\theta}$ , we could increase  $t(\theta)$  by  $\varepsilon > 0$  for all  $\theta \in [\underline{\theta}, \bar{\theta}]$ , which would preserve all incentive constraints and increase the Principal's profit, a contradiction.

**Proposition 15.7.2 (Holmstrom (1979))** *Suppose that for each  $i$ ,  $\Theta_i \subset \mathbb{R}^d$  is smoothly connected,  $v_i(y, \theta_i)$  is differentiable in  $\theta_i$  and the gradient  $D_{\theta_i} v_i(y, \theta_i)$  is uniformly bounded across  $(y, \theta_i)$ . Then any mechanism implementing an efficient decision rule  $y^*(\cdot)$  in dominant strategy is a Vickrey-Clark-Groves mechanism.*

Proof. Let  $U_i(\theta) = v_i(y^*(\theta), \theta_i) + t_i(\theta)$  denote agent  $i$ 's equilibrium utility in the mechanism. Fix some  $\underline{\theta}_i \in \Theta_i$ , and consider a smooth path  $\gamma$  from  $\underline{\theta}_i$  to  $\theta_i$ . Applying the Envelope Theorem of Milgrom and Segal (2000, Theorem 2) along path  $\gamma$ , agent  $i$ 's dominant-strategy incentive-compatibility constraints imply (as in Lemma 15.7.1):

$$U_i(\theta) = U_i(\underline{\theta}_i, \theta_{-i}) + \int_0^1 D_{\theta_i} v_i(y^*(\gamma(\tau), \theta_{-i}), \gamma(\tau)) d\gamma(\tau).$$



Let

$$S(\theta) = \sum_{i=1}^n v_i(y^*(\theta), \theta_i) = \max_{y \in X} \sum_{i=1}^n v_i(y, \theta_i).$$

Again, applying the Envelope Theorem of Milgrom and Segal (2000, Theorem 2) to this program along path  $\gamma$ , we have

$$S(\theta) = S(\underline{\theta}_i, \theta_{-i}) + \int_0^1 D_{\theta_i} v_i(y^*(\gamma(\tau), \theta_{-i}) \gamma(\tau)) d\gamma(\tau).$$

Subtracting, we get

$$U_i(\theta) - S(\theta) = U_i(\underline{\theta}_i, \theta_{-i}) - S(\underline{\theta}_i, \theta_{-i}) = d_i(\theta_{-i}),$$

and thus

$$t_i(\theta) = U_i(\theta) - v_i(y^*(\theta), \theta_i) = S(\theta) + d_i(\theta_{-i}) - v_i(y^*(\theta), \theta_i) = \sum_{j \neq i} v_j(y^*(\theta), \theta_j) + d_i(\theta_{-i}),$$

which is the transfer for a Vickrey-Clark-Groves mechanism.

The intuition behind the proof is that each agent  $i$  only has the incentive to tell the truth when he is made the residual claimant of total surplus, i.e.,  $U_i(\theta) = S(\theta) + d_i(\theta_{-i})$ , which requires a VCG mechanism. This result was also obtained by Laffont and Maskin (Econometrica, 1980), but under much more restrictive assumptions.<sup>1</sup>

### 15.7.3 Balanced VCG Mechanisms

Implementation of a surplus-maximizing decision rule  $y^*(\cdot)$ , is a necessary condition for (ex-post) Pareto efficiency, but it is not sufficient. To have Pareto efficiency, we must also guarantee that there is an ex post Balanced Budget (no waste of numeraire):

$$\sum_{i=1}^n t_i(\theta) = 0, \forall \theta \in \Theta. \quad (15.23)$$

A trivial case for which we can achieve ex post efficiency is given in the following example.

**Example 15.7.1** If there exists some agent  $i$  such that  $\Theta_i = \{\bar{\theta}_i\}$  a singleton, then we have no incentive problem for agent  $i$ ; and we can set

$$t_i(\theta) = - \sum_{j \neq i} t_j(\theta), \forall \theta \in \Theta.$$

This trivially guarantees a balanced budget.

---

<sup>1</sup>A particular case of the proposition was also obtained by Green and Laffon (1979). Namely, they allow agents to have all possible valuations over a finite decision set  $X$ , which can be described by the Euclidean type space  $\Theta_i = \mathbb{R}^{|K|}$  (the valuation of agent  $i$  for decision  $k$  being represented by  $\theta_{ik}$ ).

However, in general, there is no such a positive result. When the set of  $v(\cdot, \cdot)$  functions is "rich" then there may be no choice rule  $f(\cdot) = (y^*(\cdot), t_1(\cdot), \dots, t_I(\cdot))$  where  $y^*(\cdot)$  is ex post optimal, that is truthfully implementable in dominant, and that satisfies (15.23) so it does not result in Pareto efficient outcome.

To have an example where budget balance cannot be achieved, consider a setting with two agents. Under the assumptions of Proposition 15.7.2, any mechanism implementing an efficient decision rule  $y^*(\cdot)$  in dominant strategy is a VCG mechanism, hence

$$t_i(\theta) = v_{-i}(y^*(\theta), \theta_{-i}) + d_i(\theta_{-i}).$$

Budget balance requires that

$$0 = t_1(\theta) + t_2(\theta) = v_1(y^*(\theta), \theta_1) + v_2(y^*(\theta), \theta_2) + d_1(\theta_2) + d_2(\theta_1).$$

Letting  $S(\theta) = v_1(y^*(\theta), \theta_1) + v_2(y^*(\theta), \theta_2)$  denote the maximum surplus in state  $\theta$ , we must therefore have

$$S(\theta) = -d_1(\theta_2) - d_2(\theta_1).$$

Thus, efficiency can only be achieved when maximal total surplus is additively separable in the agents' types, which is unlikely.

To have a robust example where additive separability does not hold, consider the "public good setting" in which  $Y = [\underline{y}, \bar{y}] \subset \mathbb{R}$  with  $\underline{y} < \bar{y}$ , and for each agent  $i$ ,  $\Theta_i = [\underline{\theta}_i, \bar{\theta}_i] \subset \mathbb{R}$ ,  $\underline{\theta}_i < \bar{\theta}_i$  (to rule out the situation described in the above Example),  $v_i(y, \theta_i)$  is differentiable in  $\theta_i$ ,  $\frac{\partial v_i(y, \theta_i)}{\partial \theta_i}$  is bounded,  $v_i(y, \theta_i)$  has the single-crossing property (SCP) in  $(y, \theta_i)$  and  $y^*(\theta_1, \theta_2)$  is in the interior of  $Y$ .

Note that by the Envelope Theorem, we have (almost everywhere)

$$\frac{\partial S(\theta)}{\partial \theta_1} = \frac{\partial v_1(y^*(\theta_1, \theta_2), \theta_1)}{\partial \theta_1}.$$

It is clear that if  $y^*(\theta_1, \theta_2)$  depends on  $\theta_2$  and  $\frac{\partial v_1(y, \theta_1)}{\partial \theta_1}$  depends on  $y$ , in general  $\frac{\partial S(y, \theta)}{\partial \theta_1}$  will depend on  $\theta_2$ , and therefore  $S(\theta)$  will not be additively separable. In fact,

- SCP of  $v_2(\cdot)$  implies that  $y^*(\theta_1, \theta_2)$  is strictly increasing in  $\theta_2$ .
- SCP of  $v_1(\cdot)$  implies that  $\frac{\partial v_1(y, \theta_1)}{\partial y}$  is strictly increasing in  $\theta_1$ .

Consequently,  $\frac{\partial S(\theta)}{\partial \theta_1}$  is strictly increasing in  $\theta_2$ , thus  $S(\theta)$  is not additively separable. In this case, no ex post efficient choice rule is truthfully implementable in dominant strategies.

If the wasted money is subtracted from the social surplus, it is no longer socially efficient to implement the decision rule  $y^*(\cdot)$ . Instead, we may be interested, e.g., in maximizing the expectation

$$E \left[ \sum_i v_i(x(\theta), \theta_i) + \sum_i t_i(\theta) \right]$$

in the class of choice rules  $(y(\cdot), t_1(\cdot), \dots, t_I(\cdot))$  satisfying dominant incentive compatibility, given a probability distribution  $\varphi$  over states  $\theta$ . In this case, we may have interim efficiency as we will see in the section on Bayesian implementation.

Nevertheless, when economic environments are “thin”, we can get some positive results. Groves and Loeb (1975), Tian (1996a), and Liu and Tian (1999) provide such a class of economic environments. An example of such utility functions are:  $u_i(t_i, y) = t_i + v_i(y)$  with  $v_i(y) = -1/2y^2 + \theta_i y$  for  $i = 1, 2, \dots, n$  with  $n \geq 3$ .

## 15.8 Nash Implementation

### 15.8.1 Nash Equilibrium and General Mechanism Design

From Hurwicz’s impossibility theorem, we know that, if one wants to have a mechanism that results in Pareto efficient and individually rational allocations, one must give up the dominant strategy implementation, and then, by Revelation Principle, we must look at more general mechanisms  $\langle M, h \rangle$  instead of using a revelation mechanism.

We know the dominant strategy equilibrium is a very strong solution concept. Now, if we adopt the Nash equilibrium as a solution concept to describe individuals’ self-interested behavior, can we design an incentive compatible mechanism which implements Pareto efficient allocations?

For  $e \in E$ , a mechanism  $\langle M, h \rangle$  is said to have a *Nash equilibrium*  $m^* \in M$  if

$$h_i(m^*) \succsim_i h_i(m_i, m_{-i}^*) \quad (15.24)$$

for all  $m_i \in M_i$  and all  $i$ . Denote by  $NE(e, \Gamma)$  the set of all Nash equilibria of the mechanism  $\Gamma$  for  $e \in E$ .

It is clear every dominant strategy equilibrium is a Nash equilibrium, but the converse may not be true.

A mechanism  $\Gamma = \langle M, h \rangle$  is said to Nash-implement a social choice correspondence  $F$  on  $E$  if for every  $e \in E$

- (a)  $NE(e, \Gamma) \neq \emptyset$ ;
- (b)  $h(NE(e, \Gamma)) \subseteq F(e)$ .

It fully Nash implements a social choice correspondence  $F$  on  $E$  if for every  $e \in E$

- (a)  $NE(e, \Gamma) \neq \emptyset$
- (b)  $h(NE(e, \Gamma)) = F(e)$ .

The following proposition shows that, if a truth-telling about their characteristics is a Nash equilibrium of the revelation mechanism, it must be a dominant strategy equilibrium of the mechanism.

**Proposition 15.8.1** *For a revelation mechanism  $\Gamma = \langle E, h \rangle$ , a truth-telling  $e^*$  is a Nash equilibrium if and only if it is a dominant equilibrium*

$$h(e_i^*, e_{-i}) \succsim_i h(e_i, e_{-i}) \quad \forall (e_i, e_{-i}) \in E \& i \in N. \quad (15.25)$$

Proof. Since for every  $e \in E$  and  $i$ , by Nash equilibrium, we have

$$h(e_i, e_{-i}) \succsim_i h(e'_i, e_{-i}) \quad \text{for all } e'_i \in E_i.$$

Since this is true for any  $(e'_i, e_{-i})$ , it is a dominant strategy equilibrium. The proof is completed.

Thus, we cannot get any new results if one insists on the choice of revelation mechanisms. To get more satisfactory results, one must give up the revelation mechanism, and look for a more general mechanism with general message spaces.

Notice that, when one adopts the Nash equilibrium as a solution concept, the weak Nash implementation may not be a useful requirement. To see this, consider any social choice correspondence  $F$  and the following mechanism: each individual's message space consists of the set of economic environments, i.e., it is given by  $M_i = E$ . The outcome

function is defined as  $h(m) = a \in F(e)$  when all agents report the same economic environment  $m_i = e$ , and otherwise it is seriously punished by giving a worse outcome. Then, it is clear the truth-telling is a Nash equilibrium. However, it has a lot of Nash equilibria, in fact infinity number of Nash equilibria. Any false reporting about the economic environment  $m_i = e'$  is also a Nash equilibrium. So, when we use Nash equilibrium as a solution concept, we need a social choice rule to be implemented or full implemented in Nash equilibrium.

### 15.8.2 Characterization of Nash Implementation

Now we discuss what kind of social choice rules can be implemented through Nash incentive compatible mechanism. Maskin in 1977 gave necessary and sufficient conditions for a social choice rule to be Nash implementable (This paper was not published till 1999 due to the incorrectness of the original proof. It then appeared in *Review of Economic Studies*, 1999). Maskin's result is fundamental since it not only helps us to understand what kind of social choice correspondence can be Nash implemented, but also gives basic techniques and methods for studying if a social choice rule is implementable under other solution concepts.

Maskin's monotonicity condition can be stated in two different ways although they are equivalent.

**Definition 15.8.1 (Maskin's Monotonicity)** A social choice correspondence  $F : E \rightarrow \rightarrow A$  is said to be Maskin's monotonic if for any  $e, \bar{e} \in E$ ,  $x \in F(e)$  such that for all  $i$  and all  $y \in A$ ,  $x \succsim_i y$  implies that  $x \bar{\succsim}_i y$ , then  $x \in F(\bar{e})$ .

In words, Maskin's monotonicity requires that if an outcome  $x$  is socially optimal with respect to economy  $e$  and then economy is changed to  $\bar{e}$  so that in each individual's ordering,  $x$  does not fall below an outcome that is not below before, then  $x$  remains socially optimal with respect to  $\bar{e}$ .

**Definition 15.8.2 (Another Version of Maskin's Monotonicity)** A equivalent condition for a social choice correspondence  $F : E \rightarrow \rightarrow A$  to be Maskin's monotonic is that, if for any two economic environments  $e, \bar{e} \in E$ ,  $x \in F(e)$  such that  $x \notin F(\bar{e})$ , there is an agent  $i$  and another  $y \in A$  such that  $x \succsim_i y$  and  $y \bar{\succ}_i x$ .

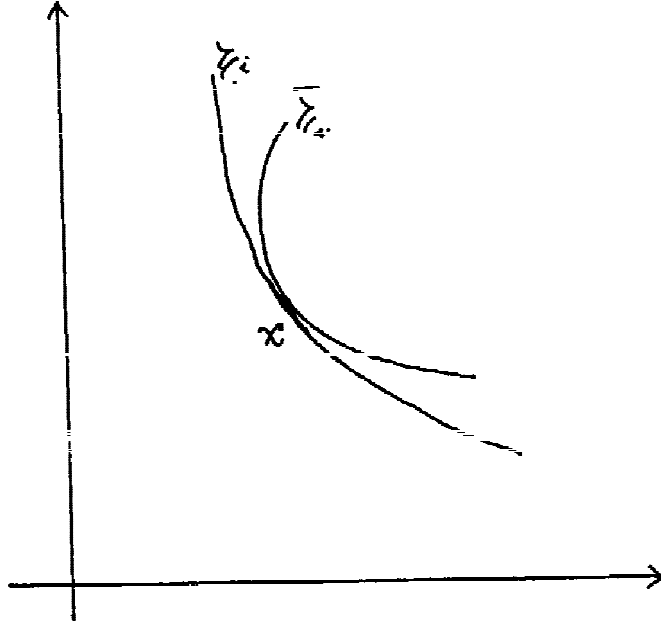


Figure 15.3: An illustration of Maskin's monotonicity.

Maskin's monotonicity is a reasonable condition, and many well known social choice rules satisfy this condition.

**Example 15.8.1 (Weak Pareto Efficiency)** The weak Pareto optimal correspondence  $P_w : E \rightarrow \rightarrow A$  is Maskin's monotonic.

Proof. If  $x \in P_w(e)$ , then for all  $y \in A$ , there exists  $i \in N$  such that  $x \succ_i y$ . Now if for any  $j \in N$  such that  $x \succ_j y$  implies  $x \bar{\succ}_j y$ , then we have  $x \bar{\succ}_i y$  for particular  $i$ . Thus,  $x \in P_w(\bar{e})$ .

**Example 15.8.2 (Majority Rule)** The majority rule or call the Condorcet correspondence  $CON : E \rightarrow \rightarrow A$  for strict preference profile, which is defined by

$$CON(e) = \{x \in A : \#\{i | x \succ_i y\} \geq \#\{i | y \succ_i x\} \text{ for all } y \in A\}$$

is Maskin's monotonic.

Proof. If  $x \in CON(e)$ , then for all  $y \in A$ ,

$$\#\{i | x \succ_i y\} \geq \#\{i | y \succ_i x\}. \quad (15.26)$$

But if  $\bar{e}$  is an economy such that, for all  $i$ ,  $x \succ_i y$  implies  $x \bar{\succ}_i y$ , then the left-hand side of (15.26) cannot fall when we replace  $e$  by  $\bar{e}$ . Furthermore, if the right-hand side rises,

then we must have  $x \succ_i y$  and  $y \bar{\succ}_i x$  for some  $i$ , a contradiction of the relation between  $e$  and  $\bar{e}$ , given the strictness of preferences. So  $x$  is still a majority winner with respect to  $\bar{e}$ , i.e.,  $x \in CON(\bar{e})$ .

In addition to the above two examples, Walrasian correspondence and Lindahl correspondence with interior allocations are Maskin's monotonic. The class of preferences that satisfy "single-crossing" property and individuals' preferences over lotteries satisfy the von Neumann-Morgenstern axioms also automatically satisfy Maskin's monotonicity.

The following theorem shows the Maskin's monotonicity is a necessary condition for Nash-implementability.

**Theorem 15.8.1** *For a social choice correspondence  $F : E \rightarrow A$ , if it is fully Nash implementable, then it must be Maskin's monotonic.*

Proof. For any two economies  $e, \bar{e} \in E$ ,  $x \in F(e)$ , then by full Nash implementability of  $F$ , there is  $m \in M$  such that  $m$  is a Nash equilibrium and  $x = h(m)$ . This means that  $h(m) \succ_i h(m'_i, m_{-i})$  for all  $i$  and  $m'_i \in M_i$ . Given  $x \succ_i y$  implies  $x \bar{\succ}_i y$ ,  $h(m) \bar{\succ}_i h(m'_i, m_{-i})$ , which means that  $m$  is also a Nash equilibrium at  $\bar{e}$ . Thus, by Nash implementability again, we have  $x \in F(\bar{e})$ .

Maskin's monotonicity itself can not guarantee a social choice correspondence is fully Nash implementable. However, under the so-called no-veto power, it becomes sufficient.

**Definition 15.8.3 (No-Veto Power)** A social choice correspondence  $F : E \rightarrow A$  is said to satisfy no-veto power if whenever for any  $i$  and  $e$  such that  $x \succ_j y$  for all  $y \in A$  and all  $j \neq i$ , then  $x \in F(e)$ .

The no-veto power (NVP) condition implies that if  $n - 1$  agents regard an outcome is the best to them, then it is social optimal. This is a rather weaker condition. NVP is satisfied by virtually all "standard" social choice rules, including weak Pareto efficient and Condorect correspondences. It is also often automatically satisfied by any social choice rules when the references are restricted. For example, for private goods economies with at least three agents, if each agent's utility function is strict monotonic, then there is no other allocation such that  $n - 1$  agents regard it best, so the no-veto power condition holds. The following theorem is given by Maskin (1977, 1999), but a complete proof of the theorem was due to Repullo (1987).

**Theorem 15.8.2** *Under no-veto power, if Maskin's monotonicity condition is satisfied, then  $F$  is fully Nash implementable.*

Proof. The proof is by construction. For each agent  $i$ , his message space is defined by

$$M_i = E \times A \times \mathcal{N}$$

where  $\mathcal{N} = \{1, 2, \dots\}$ . Its elements are denoted by  $m_i = (e^i, a^i, v^i)$ , which means each agent  $i$  announces an economic profile, an outcome, and a real number. Notice that we have used  $e^i$  and  $a^i$  to denote the economic profile of all individuals' economic characteristics and the outcome announced by individual  $i$ , but not just agent  $i$ 's economic characteristic and outcome.

The outcome function is constructed in three rules:

Rule(1). If  $m_1 = m_2 = \dots = m_n = (e, a, v)$  and  $a \in F(e)$ , the outcome function is defined by

$$h(m) = a.$$

In words, if players are unanimous in their strategy, and their proposed alternative  $a$  is  $F$ -optimal, given their proposed profile  $e$ , the outcome is  $a$ .

Rule (2). For all  $j \neq i$ ,  $m_j = (e, a, v)$ ,  $m_i = (e^i, a^i, v^i) \neq (e, a, v)$ , and  $a \in F(e)$ , define:

$$h(m) = \begin{cases} a^i & \text{if } a^i \in L(a, e_i) \\ a & \text{if } a^i \notin L(a, e_i). \end{cases}$$

where  $L(a, e_i) = \{b \in A : a R_i b\}$  which is the lower contour set of  $R_i$  at  $a$ . That is, suppose that all players but one play the same strategy and, given their proposed profile, their proposed alternative  $a$  is  $F$ -optimal. Then, the odd-man-out, gets his proposed alternative, provided that it is in the lower contour set at  $a$  of the ordering that the other players propose for him; otherwise, the outcome is  $a$ .

Rule(3). If neither Rule (1) nor Rule (2) applies, then define

$$h(m) = a^{i^*}$$

where  $i^* = \max\{i \in \mathcal{N} : v^i = \max_j v^j\}$ . In other words, when neither Rule (1) nor Rule (2) applies, the outcome is the alternative proposed by player with the highest index among those whose proposed number is maximal.



Now we show that the mechanism  $\langle M, h \rangle$  defined above fully Nash implements social choice correspondence  $F$ , i.e.,  $h(N(e)) = F(e)$  for all  $e \in E$ . We first show that  $F(e) \subset h(N(e))$  for all  $e \in E$ , i.e., we need to show that for all  $e \in E$  and  $a \in F(e)$ , there exists a  $m \in M$  such that  $a = h(m)$  is a Nash equilibrium outcome. To do so, we only need to show that any  $m$  which is given by Rule (1) is a Nash equilibrium. Note that  $h(m) = a$  and for any given  $m'_i = (e^i, a^i, v^i) \neq m_i$ , by Rule (2), we have

$$h(m'_i, m_{-i}) = \begin{cases} a^i & \text{if } a^i \in L(a, e_i) \\ a & \text{if } a^i \notin L(a, e_i). \end{cases}$$

and thus

$$h(m) R_i h(m'_i, m_{-i}) \quad \forall m'_i \in M_i.$$

Hence,  $m$  is a Nash equilibrium.

We now show that for each economic environment  $e \in E$ , if  $m$  is a Nash equilibrium, then  $h(m) \in F(e)$ . First, consider the Nash equilibrium  $m$  is given by Rule (1) so that  $a \in F(e)$ , but the true economic profile is  $e'$ , i.e.,  $m \in NE(e', \Gamma)$ . We need to show  $a \in F(e')$ . By Rule (1),  $h(m) = a$ . Let  $b \in L(a, e_i)$ , so the new message for  $i$  is  $m'_i = (e, b, v^i)$ . Then  $h(m'_i, m_{-i}) = b$  by Rule (2). Now, since  $a$  is a Nash equilibrium outcome with respect to  $e'$ , we have  $a = h(m) R'_i h(m'_i, m_{-i}) = b$ . Thus, we have shown that for all  $i \in N$  and  $b \in A$ ,  $a R_i b$  implies  $a R'_i b$ . Thus, by Maskin's monotonicity condition, we have  $a \in F(e')$ .

Next, suppose Nash equilibrium  $m$  for  $e'$  is in the Case (2), i.e, for all  $j \neq i$ ,  $m_j = (e, a, v)$ ,  $m_i \neq (e, a, v)$ . Let  $a' = h(m)$ . By Case (3), each  $j \neq i$  can induce any outcome  $b \in A$  by choosing  $(R^j, b, v^j)$  with sufficiently a large  $v^j$  (which is greater than  $\max_{k \neq j} v_k$ ), as the outcome at  $(m'_j, m_{-j})$ , i.e.,  $b = h(m'_j, m_{-j})$ . Hence,  $m$  is a Nash equilibrium with respect to  $e'$  implies that for all  $j \neq i$ , we have

$$a' R'_j b.$$

Thus, by no-veto power assumption, we have  $a' \in F(e')$ .

The same argument as the above, if  $m$  is a Nash equilibrium for  $e'$  is given by Case (3), we have  $a' \in F(e')$ . The proof is completed.

Although Maskin's monotonicity is very weak, it is violated by some social choice rules such as Solomon's famous judgement. Solomon's solution falls under Nash equilibrium

implementation, since each woman knows who is the real mother. His solution, which consisted in threatening to cut the baby in two, is not entirely foolproof. What would he have done if the impostor has had the presence of mind to scream like a real mother? Solomon's problem can be formally described by the language of mechanism design as follows.

Two women: Anne and Bets

Two economies (states):  $E = \{\alpha, \beta\}$ , where

$\alpha$ : Anne is the real mother

$\beta$ : Bets is the real mother

Solomon has three alternatives so that the feasible set is given by  $A = \{a, b, c\}$ , where

$a$ : give the baby to Anne

$b$ : give the baby to Bets

$c$ : cut the baby in half.

Solomon's desirability (social goal) is to give the baby to the real mother,

$f(\alpha) = a$  if  $\alpha$  happens

$f(\beta) = b$  if  $\beta$  happens

Preferences of Anne and Bets:

For Anne,

at state  $\alpha$ ,  $a \succ_A^\alpha b \succ_A^\alpha c$

at state  $\beta$ :  $a \succ_A^\beta c \succ_A^\beta b$

For Bets,

at state  $\alpha$ ,  $b \succ_B^\alpha c \succ_B^\alpha a$

at state  $\beta$ :  $b \succ_B^\beta a \succ_B^\beta c$

To see Solomon's solution does not work, we only need to show his social choice goal is not Maskin's monotonic. Notice that for Anne, since

$$a \succ_A^\alpha b, c,$$

$$a \succ_A^\beta b, c,$$

and  $f(\alpha) = a$ , by Maskin's monotonicity, we should have  $f(\beta) = a$ , but we actually have  $f(\beta) = b$ . So Solomon's social choice goal is not Nash implementable.

## 15.9 Better Mechanism Design

Maskin's theorem gives necessary and sufficient conditions for a social choice correspondence to be Nash implementable. However, due to the general nature of the social choice rules under consideration, the implementing mechanisms in proving characterization theorems turn out to be quite complex. Characterization results show what is possible for the implementation of a social choice rule, but not what is realistic. Thus, like most characterization results in the literature, Maskin's mechanism is not natural in the sense that it is not continuous; small variations in an agent's strategy choice may lead to large jumps in the resulting allocations, and further it has a message space of infinite dimension since it includes preference profiles as a component. In this section, we give some mechanisms that have some desired properties.

### 15.9.1 Groves-Ledyard Mechanism

Groves-Ledyard Mechanism (1977, *Econometrica*) was the first to give a specific mechanism that Nash implements Pareto efficient allocations for public goods economies.

To show the basic structure of the Groves-Ledyard mechanism, consider a simplified Groves-Ledyard mechanism. Public goods economies under consideration have one private good  $x_i$ , one public good  $y$ , and three agents ( $n = 3$ ). The production function is given by  $y = v$ .

The mechanism is defined as follows:

$M_i = R_i$ ,  $i = 1, 2, 3$ . Its elements,  $m_i$ , can be interpreted as the proposed contribution (or tax) that agent  $i$  is willing to make.

$t_i(m) = m_i^2 + 2m_j m_k$ : the actual contribution  $t_i$  determined by the mechanism with the reported  $m_i$ .

$y(m) = (m_1 + m_2 + m_3)^2$ : the level of public good  $y$ .

$x_i(m) = w_i - t_i(m)$ : the consumption of the private good.

Then the mechanism is balanced since

$$\begin{aligned} & \sum_{i=1}^3 x_i + \sum_{i=1}^3 t_i(m) \\ &= \sum_{i=1}^3 x_i + (m_1 + m_2 + m_3)^2 \\ &= \sum_{i=1}^3 x_i + y = \sum_{i=1}^3 w_i \end{aligned}$$

The payoff function is given by

$$\begin{aligned} v_i(m) &= u_i(x_i(m), y(m)) \\ &= u_i(w_i - t_i(m), y(m)). \end{aligned}$$

To find a Nash equilibrium, we set

$$\frac{\partial v_i(m)}{\partial m_i} = 0 \quad (15.27)$$

Then,

$$\frac{\partial v_i(m)}{\partial m_i} = \frac{\partial u_i}{\partial x_i}(-2m_i) + \frac{\partial u_i}{\partial y} 2(m_1 + m_2 + m_3) = 0 \quad (15.28)$$

and thus

$$\frac{\frac{\partial u_i}{\partial y}}{\frac{\partial u_i}{\partial x_i}} = \frac{m_i}{m_1 + m_2 + m_3}. \quad (15.29)$$

When  $u_i$  are quasi-concave, the first order condition will be a sufficient condition for Nash equilibrium.

Making summation, we have at Nash equilibrium

$$\sum_{i=1}^3 \frac{\frac{\partial u_i}{\partial y}}{\frac{\partial u_i}{\partial x_i}} = \sum_{i=1}^3 \frac{m_i}{m_1 + m_2 + m_3} = 1 = \frac{1}{f'(v)} \quad (15.30)$$

that is,

$$\sum_{i=1}^3 MRS_{yx_i} = MRTS_{yv}.$$

Thus, the Lindahl-Samuelson and balanced conditions hold which means every Nash equilibrium allocation is Pareto efficient.

They claimed that they have solved the free-rider problem in the presence of public goods. However, economic issues are usually complicated. Some agreed that they indeed solved the problem, some did not. There are two weakness of Groves-Ledyard Mechanism: (1) it is not individually rational: the payoff at a Nash equilibrium may be lower than at the initial endowment, and (2) it is not individually feasible:  $x_i(m) = w_i - t_i(m)$  may be negative.

How can we design the incentive mechanism to pursue Pareto efficient and individually rational allocations?

### 15.9.2 Walker's Mechanism

Walker (1981, Econometrica) gave such a mechanism. Again, consider public goods economies with  $n$  agents, one private good, and one public good, and the production function is given by  $y = f(v) = v$ . We assume  $n \geq 3$ .

The mechanism is defined by:

$$M_i = R$$

$$y(m) = \sum_{i=1}^n m_i: \text{ the level of public good.}$$

$$q_i(m) = \frac{1}{n} + m_{i+1} - m_{i+2}: \text{ personalized price of public good.}$$

$$t_i(m) = q_i(m)y(m): \text{ the contribution (tax) made by agent } i.$$

$$x_i(m) = w_i - t_i(m) = w_i - q_i(m)y(m): \text{ the private good consumption.}$$

Then, the budget constraint holds:

$$x_i(m) + q_i(m)y(m) = w_i \quad \forall m_i \in M_i \quad (15.31)$$

Making summation, we have

$$\sum_{i=1}^n x_i(m) + \sum_{i=1}^n q_i(m)y(m) = \sum_{i=1}^n w_i$$

and thus

$$\sum_{i=1}^n x_i(m) + y(m) = \sum_{i=1}^n w_i$$

which means the mechanism is balanced.

The payoff function is

$$\begin{aligned} v_i(m) &= u_i(x_i, y) \\ &= u_i(w_i - q_i(m)y(m), y(m)) \end{aligned}$$

The first order condition for interior allocations is given by

$$\begin{aligned} \frac{\partial v_i}{\partial m_i} &= -\frac{\partial u_i}{\partial x_i} \left[ \frac{\partial q_i}{\partial m_i} y(m) + q_i(m) \frac{\partial y(m)}{\partial m_i} \right] + \frac{\partial u_i}{\partial y} \frac{\partial y(m)}{\partial m_i} \\ &= -\frac{\partial u_i}{\partial x_i} q_i(m) + \frac{\partial u_i}{\partial y} = 0 \\ \Rightarrow \frac{\frac{\partial u_i}{\partial y}}{\frac{\partial u_i}{\partial x_i}} &= q_i(m) \quad (\text{FOC}) \text{ for the Lindahl Allocation} \\ \Rightarrow N(e) &\subseteq L(e) \end{aligned}$$

Thus, if  $u_i$  are quasi-concave, it is also a sufficient condition for Lindahl equilibrium. We can also show every Lindahl allocation is a Nash equilibrium allocation, i.e.,

$$L(e) \subseteq N(e) \quad (15.32)$$

Indeed, suppose  $[(x^*, y^*), q_1^*, \dots, q_n^*]$  is a Lindahl equilibrium. Let  $m^*$  be the solution of the following equation

$$\begin{aligned} q_i^* &= \frac{1}{n} + m_{i+1} - m_{i+2}, \\ y^* &= \sum_{i=1}^n m_i \end{aligned}$$

Then we have  $x_i(m^*) = x_i^*$ ,  $y(m^*) = y^*$  and  $q_i(m^*) = q_i^*$  for all  $i \in N$ . Thus from  $(x(m_i^*, m_{-i}), y(m_i^*, m_{-i})) \in \mathbb{R}_+^2$  and  $x_i(m_i^*, m_{-i}) + q_i(m^*)y(m_i^*, m_{-i}) = w_i$  for all  $i \in N$  and  $m_i \in M_i$ , we have  $(x_i(m^*), y(m^*)) R_i (x_i(m_i^*, m_{-i}), y(m_i^*, m_{-i}))$ , which means that  $m^*$  is a Nash equilibrium.

Thus, Walker's mechanism fully implements Lindahl allocations which are Pareto efficient and individually rational.

Walker's mechanism also has a disadvantage that it is not feasible although it does solve the individual rationality problem. If a person claims large amounts of  $t_i$ , consumption of private good may be negative, i.e.,  $x_i = w_i - t_i < 0$ . Tian proposed a mechanism that overcomes Walker's mechanism's problem. Tian's mechanism is individually feasible, balanced, and continuous.

### 15.9.3 Tian's Mechanism

In Tian's mechanism (JET, 1991), everything is the same as Walker's, except that  $y(m)$  is given by

$$y(m) = \begin{cases} a(m) & \text{if } \sum_{i=1}^n m_i > a(m) \\ \sum_{i=1}^n m_i & \text{if } 0 \leq \sum_{i=1}^n m_i \leq a(m) \\ 0 & \text{if } \sum_{i=1}^n m_i < 0 \end{cases} \quad (15.33)$$

where  $a(m) = \min_{i \in N'(m)} \frac{w_i}{q_i(m)}$  that can be regarded as the feasible upper bound for having a feasible allocation. Here  $N'(m) = \{i \in N : q_i(m) > 0\}$ .

An interpretation of this formulation is that if the total contributions that the agents are willing to pay were between zero and the feasible upper bound, the level of public good to be produced would be exactly the total taxes; if the total taxes were less than zero, no public good would be produced; if the total taxes exceeded the feasible upper bound, the level of the public good would be equal to the feasible upper bound.

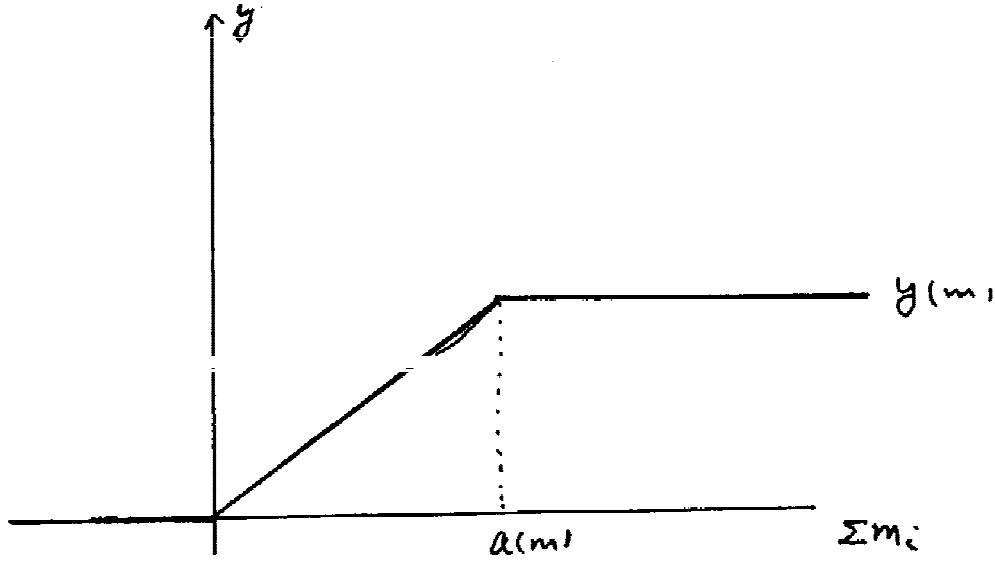


Figure 15.4: The feasible public good outcome function  $Y(m)$ .

To show this mechanism has all the nice properties, we need to assume that preferences are strictly monotonically increasing and convex, and further assume that every interior allocation is preferred to any boundary allocations: For all  $i \in N$ ,  $(x_i, y) \succ_i (x'_i, y')$  for all  $(x_i, y) \in \mathbb{R}_{++}^2$  and  $(x'_i, y') \in \partial \mathbb{R}_+^2$ , where  $\partial \mathbb{R}_+^2$  is the boundary of  $\mathbb{R}_+^2$ .

To show the equivalence between Nash allocations and Lindahl allocations. We first prove the following lemmas.

**Lemma 15.9.1** *If  $(x(m^*), y(m^*)) \in N_{M,h}(e)$ , then  $(x_i(m^*), y(m^*)) \in \mathbb{R}_{++}^2$  for all  $i \in N$ .*

Proof: We argue by contradiction. Suppose  $(x_i(m^*), y(m^*)) \in \partial \mathbb{R}_{++}^2$ . Then  $x_i(m^*) = 0$  for some  $i \in N$  or  $y(m^*) = 0$ . Consider the quadratic equation

$$y = \frac{w^*}{2(y + c)}, \quad (15.34)$$

where  $w^* = \min_{i \in N} w_i$ ;  $c = b + n \sum_{i=1}^n |m_i^*|$ , where  $b = 1/n$ . The larger root of (15.34) is positive and denoted by  $\tilde{y}$ . Suppose that player  $i$  chooses his/her message  $m_i = \tilde{y} - \sum_{j \neq i}^n m_j^*$ . Then  $\tilde{y} = m_i + \sum_{j \neq i}^n m_j^* > 0$  and

$$\begin{aligned} w_j - q_j(m_i^*, m_{-i})\tilde{y} &\geq w_j - [b + (n \sum_{s=1}^n |m_s^*| + \tilde{y})]\tilde{y} \\ &= w_j - (\tilde{y} + b + n \sum_{s=1}^n |m_s^*|)\tilde{y} \\ &= w_j - w^*/2 \geq w_j/2 > 0 \end{aligned} \quad (15.35)$$

for all  $j \in N$ . Thus,  $y(m_i^*, m_{-i}) = \tilde{y} > 0$  and  $x_j(m_i^*, m_{-i}) = w_j - q_j(m_i^*, m_{-i})y(m_i^*, m_{-i}) = w_j - q_j(m_i^*, m_{-i})\tilde{y} > 0$  for all  $j \in N$ . Thus  $(x_i(m_i^*, m_{-i}), y(m_i^*, m_{-i})) P_i (x_i(m^*), y(m^*))$  by that every interior allocation is preferred to any boundary allocations, which contradicts the hypothesis  $(x(m^*), y(m^*)) \in N_{M,h}(e)$ . Q.E.D.

**Lemma 15.9.2** *If  $(x(m^*), y(m^*)) \in N_{M,h}(e)$ , then  $y(m^*)$  is an interior point of  $[0, a(m)]$  and thus  $y(m^*) = \sum_{i=1}^n m_i^*$ .*

Proof: By Lemma 15.9.1,  $y(m^*) > 0$ . So we only need to show  $y(m^*) < a(m^*)$ . Suppose, by way of contradiction, that  $y(m^*) = a(m^*)$ . Then  $x_j(m^*) = w_j - q_j(m^*)y(m^*) = w_j - q_j(m^*)a(m^*) = w_j - w_j = 0$  for at least some  $j \in N$ . But  $x(m^*) > 0$  by Lemma 15.9.1, a contradiction. Q.E.D.

**Proposition 15.9.1** *If the mechanism has a Nash equilibrium  $m^*$ , then  $(x(m^*), y(m^*))$  is a Lindahl allocation with  $(q_1(m^*), \dots, q_n(m^*))$  as the Lindahl price vector, i.e.,  $N_{M,h}(e) \subseteq L(e)$ .*



Proof: Let  $m^*$  be a Nash equilibrium. Now we prove that  $(x(m^*), y(m^*))$  is a Lindahl allocation with  $(q_1(m^*), \dots, q_n(m^*))$  as the Lindahl price vector. Since the mechanism is feasible and  $\sum_{i=1}^n q_i(m^*) = 1$  as well as  $x_i(m^*) + q_i(m^*)y(m^*) = w_i$  for all  $i \in N$ , we only need to show that each individual is maximizing his/her preference. Suppose, by way of contradiction, that there is some  $(x_i, y) \in \mathbb{R}_+^2$  such that  $(x_i, y) P_i (X_i(m^*), Y(m^*))$  and  $x_i + q_i(m^*)y \leq w_i$ . Because of monotonicity of preferences, it will be enough to confine ourselves to the case of  $x_i + q_i(m^*)y = w_i$ . Let  $(x_{i\lambda}, y_\lambda) = (\lambda x_i + (1 - \lambda)x_i(m^*), \lambda y + (1 - \lambda)y(m^*))$ . Then by convexity of preferences we have  $(x_{i\lambda}, y_\lambda) P_i (x_i(m^*), y(m^*))$  for any  $0 < \lambda < 1$ . Also  $(x_{i\lambda}, y_\lambda) \in \mathbb{R}_+^2$  and  $x_{i\lambda} + q_i(m^*)y_\lambda = w_i$ .

Suppose that player  $i$  chooses his/her message  $m_i = y_\lambda - \sum_{j \neq i}^n m_j^*$ . Since  $y(m^*) = \sum_{j=1}^n m_j^*$  by Lemma 15.9.2,  $m_i = y_\lambda - y(m^*) + m_i^*$ . Thus as  $\lambda \rightarrow 0$ ,  $y_\lambda \rightarrow y(m^*)$ , and therefore  $m_i \rightarrow m_i^*$ . Since  $x_j(m^*) = w_j - q_j(m^*)y(m^*) > 0$  for all  $j \in N$  by Lemma 15.9.1, we have  $w_j - q_j(m_i^*, m_{-i})y_\lambda > 0$  for all  $j \in N$  as  $\lambda$  is a sufficiently small positive number. Therefore,  $y(m_i^*, m_{-i}) = y_\lambda$  and  $x_i(m_i^*, m_{-i}) = w_i - q_i(m^*)y(m_i^*, m_{-i}) = w_i - q_i(m^*)y_\lambda = x_{i\lambda}$ . From  $(x_{i\lambda}, y_\lambda) P_i (x_i(m^*), y(m^*))$ , we have  $(x_i(m_i^*, m_{-i}), y(m_i^*, m_{-i})) P_i (x_i(m^*), y(m^*))$ . This contradicts  $(x(m^*), y(m^*)) \in N_{M,h}(e)$ . Q.E.D.

**Proposition 15.9.2** *If  $(x^*, y^*)$  is a Lindahl allocation with the Lindahl price vector  $q^* = (q_1^*, \dots, q_n^*)$ , then there is a Nash equilibrium  $m^*$  of the mechanism such that  $x_i(m^*) = x_i^*$ ,  $q_i(m^*) = q_i^*$ , for all  $i \in N$ ,  $y(m^*) = y^*$ , i.e.,  $L(e) \subseteq N_{M,h}(e)$ .*

Proof: We need to show that there is a message  $m^*$  such that  $(x^*, y^*)$  is a Nash allocation. Let  $m^*$  be the solution of the following linear equations system:

$$\begin{aligned} q_i^* &= \frac{1}{n} + m_{i+1} - m_{i+2}, \\ y^* &= \sum_{i=1}^n m_i \end{aligned}$$

Then we have  $x_i(m^*) = x_i^*$ ,  $y(m^*) = y^*$  and  $q_i(m^*) = q_i^*$  for all  $i \in N$ . Thus from  $(x(m_i^*, m_{-i}), y(m_i^*, m_{-i})) \in \mathbb{R}_+^2$  and  $x_i(m_i^*, m_{-i}) + q_i(m^*)y(m_i^*, m_{-i}) = w_i$  for all  $i \in N$  and  $m_i \in M_i$ , we have  $(x_i(m^*), y(m^*)) R_i (x_i(m_i^*, m_{-i}), y(m_i^*, m_{-i}))$ . Q.E.D.

Thus, Tian's mechanism Nash implements Lindahl allocations.

## 15.10 Incomplete Information and Bayesian-Nash Implementation

### 15.10.1 Bayesian-Nash Implementation Problem

Nash implementation has imposed a strong assumption on information requirement. Although the designer does not know information about agents' characteristics, Nash equilibrium assume each agent knows characteristics of the others. This assumption is hardly satisfied in many cases in the real world. Can we remove this assumption? The answer is positive. One can use the Bayesian-Nash equilibrium, introduced by Harsanyi, as a solution concept to describe individuals' self-interested behavior. Although each agent does not know economic characteristics of the others, he knows the probability distribution of economic characteristics of the others. In this case, we can still design an incentive compatible mechanism.

Throughout we follow the notation introduced in the beginning of this chapter. Also, for simplicity, assume each individual's preferences are given by parametric utility functions  $u_i(x, \theta_i)$ . Assume that all agents and the designer know that the vector of types,  $\theta = (\theta_1, \dots, \theta_n)$  is distributed according to  $\varphi(\theta)$  a priori on a set  $\Theta$ .

Each agent knows his own type  $\theta_i$ , and therefore computes the conditional distribution of the types of the other agents:

$$\varphi(\theta_{-i}|\theta_i) = \frac{\varphi(\theta_i, \theta_{-i})}{\int_{\Theta_{-i}} \varphi(\theta_i, \theta_{-i}) d\theta_{-i}}.$$

As usual, a mechanism is a pair,  $\Gamma = \langle M, h \rangle$ . Given  $m$  with  $m_i : \Theta_i \rightarrow M_i$ , agent  $i$ 's expected utility at  $\theta_i$  is given by

$$\begin{aligned} \Pi_{\langle M, h \rangle}^i(m(\theta); \theta_i) &\equiv E_{\theta_{-i}}[u_i(h(m_i(\theta_i), m_{-i}(\theta_{-i})), \theta_i) \mid \theta_i] \\ &= \int_{\Theta_{-i}} u_i(h(m(\theta)), \theta_i) \varphi(\theta_{-i}|\theta_i) d\theta_{-i}. \end{aligned} \quad (15.36)$$

A strategy  $m^*(\cdot)$  is a *Bayesian-Nash equilibrium (BNE)* of  $\langle M, h \rangle$  if for all  $\theta_i \in \Theta_i$ ,

$$\Pi^i(m^*(\theta_i); \theta_i) \geq \Pi^i(m_i, m_{-i}^*(\theta_{-i}); \theta_i) \quad \forall m_i \in M_i.$$

That is, if player  $i$  believes that other players are using strategies  $m_{-i}^*(\cdot)$  then he maximizes his expected utility by using strategy  $m_i^*(\cdot)$ . Denote by  $B(\theta, \Gamma)$  the set of all Bayesian-Nash equilibria of the mechanism.

**Remark 15.10.1** In the present incomplete information setting, a message  $m_i : \Theta_i \rightarrow M_i$  is a dominant strategy strategy for agent  $i$  in mechanism  $\Gamma = \langle M, h \rangle$  if for all  $\theta_i \in \Theta_i$  and all possible strategies  $m_{-i}(\theta_{-i})$ ,

$$E_{\theta_{-i}}[u_i(h(m_i(\theta_i), m_{-i}(\theta_{-i})), \theta_i) \mid \theta_i] \geq E_{\theta_{-i}}[u_i(h(m'_i, m_{-i}(\theta_{-i})), \theta_i) \mid \theta_i]. \quad (15.37)$$

for all  $m'_i \in M_i$ .

Since condition (15.37) holding for all  $\theta_i$  and all  $m_{-i}(\theta_{-i})$  is equivalent to the condition that, for all  $\theta_i \in \Theta_i$ ,

$$u_i(h(m_i(\theta_i), m_{-i}), \theta_i) \geq u_i(h(m'_i, m_{-i}), \theta_i) \quad (15.38)$$

for all  $m'_i \in M_i$  and all  $m_{-i} \in M_{-i}$ ,<sup>2</sup> this leads the following definition of dominant equilibrium.

A message  $m^*$  is a dominant strategy equilibrium of a mechanism  $\Gamma = \langle M, h \rangle$  if for all  $i$  and all  $\theta_i \in \Theta_i$ ,

$$u_i(h(m_i^*(\theta_i), m_{-i}), \theta_i) \geq u_i(h(m'_i(\theta_i), m_{-i}), \theta_i)$$

for all  $m'_i \in M_i$  and all  $m_{-i} \in M_{-i}$ , which is the same as that in the case of complete information.

**Remark 15.10.2** It is clear every dominant strategy equilibrium is a Bayesian-Nash equilibrium, but the converse may not be true. Bayesian-Nash equilibrium requires more sophistication from the agents than dominant strategy equilibrium. Each agent, in order to find his optimal strategy, must have a correct prior  $\varphi(\cdot)$  over states, and must correctly predict the equilibrium strategies used by other agents.

Like Nash implementation, Bayesian-Nash incentive compatibility involves the relationship between  $F(\theta)$  and  $B(\theta, \Gamma)$ .

A mechanism  $\Gamma = \langle M, h \rangle$  is said to Bayesian-Nash implement a social choice correspondence  $F$  on  $\Theta$  if for every  $\theta \in \Theta$ ,

$$(a) \ B(\theta, \Gamma) \neq \emptyset;$$

$$(b) \ h(B(\theta, \Gamma)) \subseteq F(\theta).$$

---

<sup>2</sup>(15.37) implies (15.38) by noting  $m_{-i}(\theta_{-i}) = m_{-i}$ , and (15.38) implies (15.37) because  $E_{\theta_{-i}}[u_i(h(m_i, m_{-i}(\theta_{-i})), \theta_i) \mid \theta_i] = u_i(h(m_i, m_{-i}), \theta_i)$  for all  $m_i \in M_i$ .

It fully Bayesian-Nash implements a social choice correspondence  $F$  on  $\Theta$  if for every  $\theta \in \Theta$ ,

- (a)  $B(\theta, \Gamma) \neq \emptyset$
- (b)  $h(B(\theta, \Gamma)) = F(\theta)$ .

If for a given social choice correspondence  $F$ , there is some mechanism that Bayesian-Nash implements  $F$ , we call this social choice correspondence Bayesian-Nash implementable.

For a general social choice correspondence  $F : \Theta \rightarrow Z$ , Pastlewaite–Schmeidler (JET, 1986), Palfrey-Srivastava (RES, 1989), Mookherjee-Reichelstein (RES, 1990), Jackson (Econometrica, 1991), Dutta-Sen (Econometrica, 1994), Tian (Social Choices and Welfare, 1996; Journal of Math. Econ., 1999) provided necessary and sufficient conditions for a social choice correspondence  $F$  to be Bayesian-Nash implementable. Under some technical conditions, they showed that a social choice correspondence  $F$  is Bayesian-Nash implementable if and only if  $F$  is Bayesian monotonic and Bayesian-Nash incentive compatible.

To see specifically what is implementable in BNE, in the remainder of this section, the social choice goal is given by a single-valued social choice function  $f : \Theta \rightarrow Z$ . Furthermore, by Revelation Principle below, without loss of generality, we can focus on direct revelation mechanisms.

**Definition 15.10.1** *Choice rule  $f(\cdot)$  is truthfully implementable in BNE if for all  $\theta_i \in \Theta_i$  and  $i \in I$ ,  $m_i^*(\theta_i) = \theta_i$  is a BNE of the direct revelation mechanism  $\Gamma = (\Theta_1, \dots, \Theta_i, f(\cdot))$ , i.e.,  $\forall i, \forall \theta_i \in \Theta_i$ , we have*

$$E_{\theta_{-i}}[u_i(f(\theta_i, \theta_{-i}), \theta_i) \mid \theta_i] \geq E_{\theta_{-i}}[u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i) \mid \theta_i], \forall \hat{\theta}_i \in \Theta_i.$$

The concept of Bayesian incentive-compatibility means that every agent will report his type truthfully provided *all* other agents are employing their truthful strategies and thus every truthful strategy profile is a Bayesian equilibrium of the direct mechanism  $\langle S, x \rangle$ . Notice that Bayesian incentive compatibility does not say what is the best response of an agent when other agents are not using truthful strategies. So it may also contain some undesirable equilibrium outcomes when a mechanism has multiple equilibria. The goal for designing a mechanism is to reach a desirable equilibrium outcome, but it may also result

in an undesirable equilibrium outcome. Thus, while the incentive compatibility requirement is central, it may not be sufficient for a mechanism to give all of desirable outcomes. The severity of this multiple equilibrium problem has been exemplified by Demski and Sappington (JET, 1984), Postlewaite and Schmeidler (JET, 1986), Repullo (Econometrica, 1988), and others. The implementation problem involves designing mechanisms to ensure that all equilibria result in desirable outcomes which are captured by the social choice set.

These inequalities are called Bayesian Incentive-Compatibility Constraints.

**Proposition 15.10.1 (Revelation Principle)** *Choice rule  $f(\cdot)$  is implementable in BNE if and only if it is truthfully implementable in BNE.*

*Proof.* The proof is the same as before: Suppose that there exists a mechanism  $\Gamma = (M_1, \dots, M_n, g(\cdot))$  and an equilibrium strategy profile  $m^*(\cdot) = (m_1^*(\cdot), \dots, m_n^*(\cdot))$  such that  $g(m^*(\cdot)) = f(\cdot)$  and  $\forall i, \forall \theta_i \in \Theta_i$ , there is

$$E_{\theta_{-i}}[u_i(g(m_i^*(\theta_i), m_{-i}^*(\theta_{-i})), \theta_i) \mid \theta_i] \geq E_{\theta_{-i}}[u_i(g(m'_i, m_{-i}^*(\theta_{-i})), \theta_i) \mid \theta_i], \forall m'_i \in M_i.$$

One way to deviate for agent  $i$  is by pretending that his type is  $\hat{\theta}_i$  rather than  $\theta_i$ , i.e., sending message  $m'_i = m_i^*(\hat{\theta}_i)$ . This gives

$$E_{\theta_{-i}}[u_i(g(m_i^*(\theta_i), m_{-i}^*(\theta_{-i})), \theta_i) \mid \theta_i] \geq E_{\theta_{-i}}[u_i(g(m_i^*(\hat{\theta}_i), m_{-i}^*(\theta_{-i})), \theta_i) \mid \theta_i], \forall \hat{\theta}_i \in \Theta_i.$$

But since  $g(m^*(\theta)) = f(\theta)$ ,  $\forall \theta \in \Theta$ , we must have  $\forall i, \forall \theta_i \in \Theta_i$ , there is

$$E_{\theta_{-i}}[u_i(f(\theta_i, \theta_{-i}), \theta_i) \mid \theta_i] \geq E_{\theta_{-i}}[u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i) \mid \theta_i], \forall \hat{\theta}_i \in \Theta_i.$$

## 15.10.2 Ex-Post Efficient Implementation by the Expected Externality Mechanism

Once again, let us return to the quasilinear setting. From the section on dominant strategy implementation, we know that there is, in general, no ex-post Pareto efficient implementation if the dominant equilibrium solution behavior is assumed. However, in quasilinear environments, relaxation of the equilibrium concept from DS to BNE allows

us to implement ex post Pareto efficient choice rule  $f(\cdot) = (y^*(\cdot), t_1(\cdot), \dots, t_i(\cdot))$ , where  $\forall \theta \in \Theta$ ,

$$y^*(\theta) \in \operatorname{argmax}_{y \in Y} \sum_{i=1}^n v_i(y, \theta_i),$$

and there is a balanced budget:

$$\sum_{i=1}^n t_i(\theta) = 0.$$

The Expected Externality Mechanism described below was suggested independently by D'Aspremont and Gerard-Varet (1979) and Arrow (1979) and are also called AGV mechanisms in the literature. This mechanism enables us to have ex-post Pareto efficient implementation under the following additional assumption.

**Assumption 15.10.1** *Types are distributed independently:  $\phi(\theta) = \Pi_i \phi_i(\theta_i), \forall \theta \in \Theta$ .*

To see this, take the VCG transfer for agent  $i$  and instead of using other agents' announced types, take the expectation over their possible types

$$t_i(\hat{\theta}) = E_{\theta_{-i}} \left[ \sum_{j \neq i} v_j(y^*(\hat{\theta}_i, \theta_{-i}), \theta_j) \right] + d_i(\hat{\theta}_{-i})$$

(by the above assumption the expectation over  $\theta_{-i}$  does not have to be taken conditionally on  $\theta_i$ ). Note that unlike in VCG mechanisms, the first term only depends on agent  $i$ 's announcement  $\hat{\theta}_i$ , and not on other agents' announcements. This is because it sums the expected utilities of agents  $j \neq i$  assuming that they tell the truth and given that  $i$  announced  $\hat{\theta}_i$ , and does not depend on the actual announcements of agents  $j \neq i$ . This means that  $t_i(\cdot)$  is less "variable", but on average it will cause  $i$ 's incentives to be lined up with the social welfare.

To see that  $i$ 's incentive compatibility is satisfied given that agent  $j \neq i$  announce truthfully, observe that agent  $i$  solves

$$\max_{\hat{\theta}_i \in \Theta_i} E_{\theta_{-i}} [v_i(y^*(\hat{\theta}_i, \theta_{-i}), \theta_j) + t_i(\hat{\theta}_i, \theta_{-i})] = \max_{\hat{\theta}_i \in \Theta_i} E_{\theta_{-i}} \left[ \sum_{j=1}^n v_j(y^*(\hat{\theta}_i, \theta_{-i}), \theta_j) \right] + d_i(\theta_{-i}).$$

Again, agent  $i$ 's announcement only matters through the decision  $y^*(\hat{\theta}_i, \theta_{-i})$ . Furthermore, by the definition of the efficient decision rule  $y^*(\cdot)$ , for each realization of  $\theta_{-i} \in \Theta_{-i}$ , agent  $i$ 's expected utility is maximized by choosing the decision  $y^*(\theta_i, \theta_{-i})$ , which can be achieved by announcing truthfully:  $\hat{\theta}_i = \theta_i$ . Therefore, truthful announcement maximizes the agent  $i$ 's expected utility as well. Thus, BIC is satisfied.

**Remark 15.10.3** The argument relies on the assumption that other agents announce truthfully. Thus, it is in general not a dominant strategy for agent  $i$  to announce the truth. Indeed, if agent  $i$  expects the other agents to lie, i.e., announce  $\hat{\theta}_{-i}(\theta_{-i}) \neq \theta_{-i}$ , then agent  $i$ 's expected utility is

$$E_{\theta_{-i}}[v_i(y^*(\hat{\theta}_i, \hat{\theta}_{-i}(\theta_{-i})), \theta_j) + \sum_{j \neq i}^n v_j(y^*(\hat{\theta}_i, \theta_{-i}), \theta_j)] + d_i(\theta_{-i}),$$

which may not be maximized by truthfully announcement.

Furthermore, we can now choose functions  $d_i(\cdot)$  so that the budget is balanced. To see this, let

$$\xi_i(\theta_i) = E_{\theta_{-i}}[\sum_{j \neq i} v_j(y^*(\theta_i, \theta_{-i}), \theta_j)],$$

so that the transfers in the Expected Externality mechanism are  $t_i(\theta) = \xi_i(\theta_i) + d_i(\theta_{-i})$ .

We will show that we can use the  $d(\cdot)$  functions to "finance" the  $\xi(\cdot)$  functions in the following way:

Let

$$d_j(\theta_{-j}) = - \sum_{i \neq j} \frac{1}{n-1} \xi_i(\theta_i).$$

Then

$$\begin{aligned} \sum_{j=1}^n d_j(\theta_{-j}) &= -\frac{1}{n-1} \sum_{j=1}^n \sum_{i \neq j} \xi_i(\theta_i) = -\frac{1}{n-1} \sum_{j \neq i} \sum_{i=1}^n \xi_i(\theta_i) \\ &= -\frac{1}{n-1} \sum_{i=1}^n (n-1) \xi_i(\theta_i) = - \sum_{i=1}^n \xi_i(\theta_i), \end{aligned}$$

and therefore

$$\sum_{i=1}^n t_i(\theta) = \sum_{i=1}^n \xi_i(\theta_i) + \sum_{i=1}^n d_i(\theta_{-i}) = 0.$$

Thus, we have shown that when agents's Bernoulli utility functions are quasilinear and agents' types are statistically independent, there is an ex-post efficient social choice function that is implementable in Bayesian-Nash equilibrium.

## Application to a Linear Model

Consider the quasi-linear environment with the decision set  $Y \subset \mathbb{R}^n$ , and the type spaces  $\Theta_i = [\underline{\theta}_i; \bar{\theta}_i] \subset \mathbb{R}$  for all  $i$ . Each agent  $i$ 's utility takes the form

$$\theta_i y_i + t_i.$$

Thus, agent  $i$  cares only about component  $y_i$  of the public outcome. Note that these payoffs satisfy SCP.

The decision set  $Y$  depends on the application.

Two examples:

1. Allocating a private good:  $Z = \{y \in \{0, 1\}^n : \sum_i y_i = 1\}$ .
2. Provision of a nonexcludable public good:  $Y = \{(q, \dots, q) \in \mathbb{R}^n : q \in \{0, 1\}\}$ .

We can fully characterize both dominant strategy (DIC) and Bayesian Incentive Compatibility (BIC) choice rules in this environment using the one-agent characterization in Lemma 15.7.1.

Let  $U_i(\theta) = \theta_i y(\theta) + t_i(\theta)$ . Applying Lemma 15.7.1 to characterize DIC constraints for any fixed  $\theta_{-i}$ , we have

**Proposition 15.10.2** *In the linear model, choice rule  $(y(\cdot), t_1(\cdot), \dots, t_I(\cdot))$  is Dominant Incentive Compatible if and only if for all  $i \in I$ ,*

- (1)  $y_i(\theta_i, \theta_{-i})$  is nondecreasing in  $\theta_i$ , (DM);
- (2)  $U_i(\theta_i, \theta_{-i}) = U_i(\underline{\theta}_i, \theta_{-i}) + \int_{\underline{\theta}_i}^{\bar{\theta}_i} y_i(\tau, \theta_{-i}) d\tau, \forall \theta \in \Theta$ . (DICFOC)

As for agent  $i$ 's interim expected utility when he announces type  $\hat{\theta}_i$ , it is

$$\Psi(\hat{\theta}_i, \theta_{-i}) = \theta_i E_{\theta_{-i}} y_i(\hat{\theta}_i, \theta_{-i}) + E_{\theta_{-i}} t_i(\hat{\theta}_i, \theta_{-i}),$$

thus, it depends only on the interim expected consumption  $E_{\theta_{-i}} y_i(\hat{\theta}_i, \theta_{-i})$  and transfer  $E_{\theta_{-i}} t_i(\hat{\theta}_i, \theta_{-i})$ . BIC means that truth-telling is optimal at the interim stage. Therefore, BIC choice rules can be fully characterized using the one-agent result in Lemma 15.7.1.

**Proposition 15.10.3** *In the linear model, choice rule  $(y(\cdot), t_1(\cdot), \dots, t_I(\cdot))$  is Bayesian Incentive Compatible if and only if for all  $i \in I$ ,*

- (1)  $E_{\theta_{-i}} y_i(\theta_i, \theta_{-i})$  is nondecreasing in  $\theta_i$  (BM);
- (2)  $E_{\theta_{-i}} U_i(\theta_i, \theta_{-i}) = E_{\theta_{-i}} U_i(\underline{\theta}_i, \theta_{-i}) + \int_{\underline{\theta}_i}^{\bar{\theta}_i} E_{\theta_{-i}} y_i(\tau, \theta_{-i}) d\tau, \forall \theta_i \in \Theta_i$ . (BIC-FOC)



Note that (BICFOC) is a pooling of (DICFOC), and similarly (BM) is a pooling of (DM). The latter means that BIC allows to implement some decision rules that are not DS implementable.

The proposition implies that in any two BIC mechanisms implementing the same decision rule  $y(\cdot)$ , the interim expected utilities  $E_{\theta_{-i}}U_i(\theta)$  and thus the transfers  $E_{\theta_{-i}}t_i(\theta)$  coincide up to a constant.

**Corollary 15.10.1** *In the linear model, for any BIC mechanism implementing the ex post Pareto efficient decision rule  $y(\cdot)$ , there exists a VCG mechanism with the same interim expected transfers and utilities.*

Proof. If  $(y^*(\cdot), t(\cdot))$  is a BIC mechanism and  $(y^*(\cdot), \tilde{t}(\cdot))$  is a VCG mechanism, then by Proposition 15.10.3,  $E_{\theta_{-i}}t_i(\theta) = E_{\theta_{-i}}\tilde{t}_i(\theta) + c_i, \forall \theta_i \in \Theta_i$ . Then letting  $\bar{t}_i(\theta) = \tilde{t}_i(\theta) + c_i$ ,  $(y^*(\cdot), \bar{t}(\cdot))$  is also a VCG mechanism, and  $E_{\theta_{-i}}\bar{t}_i(\theta) = E_{\theta_{-i}}t_i(\theta)$ .

For example, the expected externality mechanism is interim-equivalent to a VCG mechanism. Its only advantage is that it allows to balance the budget ex post, i.e., in each state of the world. More generally, if a decision rule  $y(\cdot)$  is DS implementable, the only reason to implement it in a BIC mechanism that is not DIC is whether we care about ex post transfers/utilities rather than just their interim or ex ante expectations.

### 15.10.3 Participation Constraints

Thinking of the mechanism as a contract raises the following issues:

- Will the agents accept it voluntarily, i.e., are their participation constraints satisfied?
- If one of the agents designs the contract, he will try to maximize his own payoff subject to the other agents  $i$  participation constraints. What will the optimal contract look like?

To analyze these questions, we need to impose additional restrictions on the choice rule in the form of participation constraints. These constraints depend on when agents can withdraw from the mechanism, and what they get when they do. Let  $\hat{u}_i(\theta_i)$  be the utility of agent  $i$  if he withdraws from the mechanism. (This assumes that when an agent withdraws from the mechanism he does not care what the mechanism does with other agents.)

**Definition 15.10.2** The choice rule  $f(\cdot)$  is

1. Ex Post Individually Rational if for all  $i$ ,

$$U_i(\theta) \equiv u_i(f(\theta), \theta) \geq \hat{u}_i(\theta_i), \forall \theta \in \Theta.$$

2. Interim Individually Rational if for all  $i$ ,

$$E_{\theta_{-i}}[U_i(\theta_i, \theta_{-i})] \geq \hat{u}_i(\theta_i), \forall \theta \in \Theta.$$

3. Ex ante Individually Rational if for all  $i$ ,

$$E_{\theta}[U_i(\theta)] \geq E_{\theta_i}[\hat{u}_i(\theta_i)].$$

Note that ex post IRs imply interim IR, which in turn imply ex ante IRs, but the reverse may not be true. Then, the constraints imposed by voluntary participation are most sever when agents can withdraw at the ex post stage.

Thus, when agents' types are not observed, the set of social choice functions that can be successfully implemented are those that satisfy not only the incentive compatibility, respectively, either a dominant strategy or Bayesian Nash dominant strategy, depending on the equilibrium concept used, but also participation constraints that are relevant in the environment under study.

The ex post IRs arise when the agent can withdraw in any state after the announcement of the outcome. For example, they are satisfied by any decentralized bargaining procedure. These are the hardest constraints to satisfy.

The Interim IRs arise when the agent can withdraw after learning his type  $\theta_i$ ; but before learning anything about other agents' types. Once the agent decides to participate, the outcome can be imposed on him. These constraints are easier to satisfy than ex post IR.

Finally, with the ex-ante participation constraint the agent can commit to participating even before his type is realized. These are the easiest constraints to satisfy. For example, in a quasi-linear environment, whenever the mechanism generates a positive expected surplus, i.e.,

$$E_{\theta}[\sum_i U_i(\theta)] \geq E_{\theta}[\sum_i \hat{u}_i(\theta_i)],$$

all agents' ex ante IR can be satisfied by reallocating expected surplus among agents through lump-sum transfers, which will not disturb agents' incentive constraints or budget balance.

For this reason, we will focus mainly on interim IR. In the following, we will illustrate further the limitations on the set of implementable social choice functions that may be caused by participation constraints by the important theorem given by Myerson-Satterthwaite (1983).

## The Myerson-Satterthwaite Theorem

Even though Pareto efficient mechanisms do exist (e.g. the expected externality mechanism), it remains unclear whether such a mechanism may result from private contracting among the parties. We have already seen that private contracting need not yield efficiency. In the Principal-Agent model, the Principal offers an efficient contract to extract the agent's information rent. However, this leaves open the question of whether there is some contracting/bargaining procedure that would yield efficiency. For example, in the P-A model, if the agent makes an offer to the principal, who has no private information on his own, the agent would extract all the surplus and implement efficient as a result. Therefore, we focus on a bilateral situation in which both parties have private information. In this situation, it turns out that generally there does not exist an efficient mechanism that satisfies both agents' participation constraints.

Consider the setting of allocating an indivisible object with two agents - a seller and a buyer:  $I = \{S, B\}$ . Each agent's type is  $\theta_i \in \Theta_i = [\underline{\theta}_i, \bar{\theta}_i] \subset \mathbb{R}$ , where  $\theta_i \sim \varphi_i(\cdot)$  are independent, and  $\varphi_i(\cdot) > 0$  for all  $\theta_i \in \Theta_i$ . Let  $y \in \{0, 1\}$  indicate whether  $B$  receives the good. A choice rule is then  $f(\theta) = (y(\theta), t_1(\theta), t_2(\theta))$ . The agents' utilities can then be written as

$$\begin{aligned} u_B(y, \theta_B) &= \theta_B y + t_B, \\ u_S(y, \theta_S) &= -\theta_S y + t_S. \end{aligned}$$

It is easy to see that an efficient decision rule  $y^*(\theta)$  must have<sup>3</sup>

$$y^*(\theta_B, \theta_S) = \begin{cases} 1 & \text{if } \theta_B > \theta_S, \\ 0 & \text{if } \theta_B < \theta_S. \end{cases}$$

We could use an expected externality mechanism to implement an efficient decision rule in BNE with ex post budget balance. However suppose that we have to satisfy Interim IR:

$$\begin{aligned} E_{\theta_S}[\theta_B y(\theta_B, \theta_S) + t_B(\theta_B, \theta_S)] &\geq 0, \\ E_{\theta_B}[-\theta_S y(\theta_B, \theta_S) + t_S(\theta_B, \theta_S)] &\geq 0. \end{aligned}$$

As for budget balance, let us relax this requirement by requiring only ex ante Budget Balance:

$$E_{\theta}[t_B(\theta_B, \theta_S) + t_S(\theta_B, \theta_S)] \leq 0.$$

Unlike ex post budget balance considered before, ex ante budget balance allows us to borrow money as long as we break even on average. It also allows us to have a surplus of funds. The only constraint is that we cannot have an expected shortage of funds.

We can then formulate a negative result:

**Theorem 15.10.1 (Myerson-Satterthwaite)** *In the two-party trade setting above with  $(\underline{\theta}_B, \bar{\theta}_B) \cap (\underline{\theta}_S, \bar{\theta}_S) \neq \emptyset$  (gains from trade are possible but not certain) there is no BIC choice rule that has the efficient decision rule and satisfies ex ante Budget Balance and interim IR.*

Proof. Consider first the case where  $[\underline{\theta}_B, \bar{\theta}_B] = [\underline{\theta}_S, \bar{\theta}_S] = [\underline{\theta}, \bar{\theta}]$ .

By Corollary 15.10.1 above we can restrict attention to VCG mechanisms while preserving ex ante budget balance and interim IR. Such mechanisms take the form

$$\begin{aligned} t_B(\theta_B, \theta_S) &= -\theta_S y^*(\theta_B, \theta_S) + d_B(\theta_S), \\ t_S(\theta_B, \theta_S) &= \theta_B y^*(\theta_B, \theta_S) + d_S(\theta_S). \end{aligned}$$

By interim IR of B's type  $\underline{\theta}$ , using the fact that  $y^*(\underline{\theta}, \theta_S) = 0$  with probability 1, we must have  $E_{\theta_S} d_B(\theta_S) \geq 0$ . Similarly, by interim IR of S's type  $\bar{\theta}$ , using the fact

---

<sup>3</sup>Given our full support assumption on the distributions, ex ante efficiency dictates that the decision rule coincide with  $y^*(\cdot)$  almost everywhere.

that  $y^*(\theta_B, \bar{\theta}) = 0$  with probability 1, we must have  $E_{\theta_B} d_S(\theta_B) \geq 0$ . Thus, adding the transfers, we have

$$\begin{aligned} E_{\theta}[t_B(\theta_B, \theta_S) + t_S(\theta_B, \theta_S)] &= E_{\theta}[(\theta_B - \theta_S)y^*(\theta_B, \theta_S)] + E_{\theta_S}[d_B(\theta_S)] + E_{\theta_B}[d_S(\theta_B)] \\ &\geq E_{\theta}[\max\{\theta_B - \theta_S\}] > 0 \end{aligned}$$

since  $\Pr(\theta_B > \theta_S) > 0$ . Therefore, ex ante budget balance cannot be satisfied.

Now, in the general case, let  $(\underline{\theta}, \bar{\theta}) = (\underline{\theta}_B, \bar{\theta}_B) \cap (\underline{\theta}_S, \bar{\theta}_S)$ , and observe that any type of either agent above  $\bar{\theta}$  [or below  $\underline{\theta}$ ] has the same decision rule, and therefore (by IC) must have the same transfer, as this agent's type  $\bar{\theta}$  [resp.  $\underline{\theta}$ ]. Therefore, the payments are the same as if both agents having valuations distributed on  $[\underline{\theta}, \bar{\theta}]$ ; with possible atoms on  $\underline{\theta}$  and  $\bar{\theta}$ . The argument thus still applies.

The intuition for the proof is simple: In a VCG mechanism, in order to induce truthful revelation, each agent must become the residual claimant for the total surplus. This means that in case trade is implemented, the buyer pays the seller's cost for the object, and the seller receives the buyer's valuation for the object. Any additional payments to the agents must be nonnegative, in order to satisfy interim IRs of the lowest-valuation buyer and the highest-cost seller. Thus, each agent's utility must be at least equal to the total surplus. In BNE implementation, agents receive the same expected utilities as in the VCG mechanism, thus again each agent's expected utility must equal at least to the total expected surplus. This cannot be done without having an expected infusion of funds equal to the expected surplus.

**Interpretation of the Theorem:** Who is designing the mechanism to maximize expected surplus subject to agents' interim IRs?

- Agents themselves at the ex ante stage? They would face ex ante rather than interim IRs, which would be easy to satisfy.
- An agent at the interim stage? He would be interested not in efficiency but in maximizing his own payoff.
- A benevolent mediator at the interim stage? But where would such a mediator come from?

A better interpretation of the result is as an upper bound on the efficiency of decentralized bargaining procedures. Indeed, any such procedure can be thought of as a mechanism, and must satisfy interim IR (indeed, even ex post IR), and ex ante budget balance (indeed, even ex post budget balance). The Theorem says that decentralized bargaining in this case cannot be efficient. In the terminology of the Coase Theorem, private information creates a "transaction cost".

Cramton-Gibbons-Klemperer (1987) show that when the object is divisible, and is jointly owned initially, efficiency can be attained if initial ownership is sufficiently evenly allocated. For example, suppose that the buyer initially owns  $\hat{y} \in [0, 1]$  of the object. Efficiency can be achieved when  $\hat{y}$  is sufficiently close to  $1/2$ . Thus, when a partnership is formed, the partners can choose initial shares so as to eliminate inefficiencies in dissolution. While this result can be applied to study optimal initial allocations of property rights, it does not explain why property rights are good at all. That is, interim IRs stemming from property rights can only hurt the parties in the model. For example, the parties could achieve full efficiency by writing an ex ante contract specifying the AGV mechanism and not allowing withdrawal at the interim stage. One would have to appeal to a difficulty of specify in complicated mechanisms such as AGV to explain why the parties would look for optimal property rights that facilitate efficient renegotiation rather than specifying an efficient mechanism ex ante without allow in subsequent withdrawal or renegotiation.

#### 15.10.4 The Revenue Equivalence Theorem in Auctions

Let us consider again the setting of allocating an indivisible object among  $n$  risk-neutral buyers:  $Y = \{y \in \{0, 1\}^n : \sum_i y_i = 1\}$ , and the payoff of buyer  $i$  is

$$\theta_i y_i + t_i.$$

The buyers' valuations are independently drawn from  $[\underline{\theta}_i, \bar{\theta}_i]$  with  $\underline{\theta}_i < \bar{\theta}_i$  according to a strictly positive density  $\varphi_i(\cdot)$ , and c.d.f. denoted by  $\Phi_i(\cdot)$ .

Suppose that the object initially belongs to a seller (auctioneer), who is the residual sink of payments. The auctioneer's expected revenue in a choice rule  $(y(\cdot), t_1, \dots, t_n)$  can be written as

$$-\sum_i E_{\theta} t_i(\theta) = \sum_i E_{\theta} [\theta_i y_i(\theta) - U_i(\theta)] = \sum_i E_{\theta} [\theta_i y_i(\theta)] - \sum_i E_{\theta_i} E_{\theta_{-i}} [U_i(\theta_i, \theta_{-i})].$$

The first term is the agents' expected total surplus, while the second term is the sum of the agents' expected utilities. By Proposition 15.10.3, the second term is fully determined by the decision (object allocation) rule  $y(\cdot)$  together with the lowest types' interim expected utilities  $E_{\theta_{-i}}[U_i(\underline{\theta}_i, \theta_{-i})]$ . Since the total surplus is pinned down as well, we have the following result.

**Theorem 15.10.2 (The Revenue Equivalence Theorem)** *Suppose that two different auction mechanisms have Bayesian-Nash Equilibria in which (i) the same decision (object allocation) rule  $y(\cdot)$  is implemented, and (ii) each buyer  $i$  has the same interim expected utility when his valuation is  $\underline{\theta}_i$ . Then the two equilibria of the two auction mechanisms generate the same expected revenue for the seller.*

Note that even when the decision rule and lowest agents' utilities are fixed, the seller still has significant freedom in designing the auction procedure, since there are many ways to achieve given interim expected transfers  $\bar{t}_i(\theta_i)$  with different ex post transfer  $t_i(\theta)$ .

For example, suppose that the buyers are symmetric, and that the seller wants to implement an efficient decision rule  $y^*(\cdot)$ , and make sure that the lowest-valuation buyers receive zero expected utility. We already know that this can be done in dominant strategies - using the Vickrey (second-price sealed-bid) auction. More generally, consider a  $k^{th}$  price sealed-bid auction, with  $1 \leq k \leq n$ . Here the winner is the highest bidder, but he pays the  $k^{th}$  highest bid. Suppose that buyers' valuations are i.i.d. Then it can be shown that the auction has a unique equilibrium, which is symmetric, and in which an agent's bid is an increasing function of his valuation,  $b(\theta_i)$ . (See, e.g., Fudenberg- Tirole "Game Theory," pp. 223-225.) Agent  $i$  receives the object when he submits the maximum bid, which happens when he has the maximum valuation. Therefore, the auction implements an efficient decision rule. Also, a buyer with valuation  $\underline{\theta}$  wins with zero probability, hence he has zero expected utility. Hence, the Revenue Equivalence Theorem establishes that  $k^{th}$  price sealed-bid auctions generate the same expected revenue for the seller for all  $k$ .

How can this be true, if for any given bid profile, the seller receives a higher revenue when  $k$  is lower? The answer is that the bidders will submit lower bids when  $k$  is lower, which exactly offsets the first effect. For example, we know that in the second-price auction, buyer will bid their true valuation. In the first price auction, on the other hand, buyers will obviously bid less than their true valuation, since bidding their true valuation

would give them zero expected utility. The revenue equivalence theorem establishes that the expected revenue ends up being the same. Remarkably, we know this without solving for the actual equilibrium of the auction. With  $k > 2$ , the revenue equivalence theorem implies that buyers will bid more than their true valuation.

Now we can also solve for the seller's optimal auction. Suppose the seller also has the option of keeping the object to herself. The seller will be called "agent zero", her valuation for the object denoted by  $\theta_0$ , and whether she keeps the object will be denoted by  $y_0 \in \{0, 1\}$ . Thus, we must have  $\sum_{i=0}^n y_i = 1$ .

The seller's expected payoff can be written as

$$\theta_0 E_\theta y_0(\theta) + \text{Expected Revenue} = \theta_0 E_\theta y_0(\theta) + \sum_{i=1}^n E_\theta [\theta_i y_i(\theta)] - \sum_{i=1}^n E_{\theta_i} E_{\theta_{-i}} [U_i(\theta_i, \theta_{-i})].$$

Thus, the seller's expected payoff is the difference between total surplus and the agents' expected information rents.

By Proposition 15.10.3, the buyer's expected interim utilities must satisfy

$$E_{\theta_{-i}} [U_i(\theta_i, \theta_{-i})] = E_{\theta_{-i}} [U_i(\underline{\theta}_i, \theta_{-i})] + \int_{\underline{\theta}_i}^{\bar{\theta}_i} E_{\theta_{-i}} y_i(\tau, \theta_{-i}) d\tau, \forall \theta_i \in \Theta_i. \quad (BICFOC)$$

By the buyers' interim IR,  $E_{\theta_{-i}} [U_i(\underline{\theta}_i, \theta_{-i})] \geq 0, \forall i$ , and for a given decision rule  $y$ , the seller will optimally chose transfers to set  $E_{\theta_{-i}} [U_i(\underline{\theta}_i, \theta_{-i})] = 0, \forall i$  (due to SCP, the other types' IR will then be satisfied as well - see Lemma 15.7.1 above.) Furthermore, just as in the one-agent case, upon integration by parts, buyer  $i$ 's expected information rent can be written as

$$E_\theta E_{\theta_{-i}} \left[ \frac{1}{h_i(\theta_i)} y_i(\theta) \right],$$

where  $h_i(\theta_i) = \frac{\varphi_i(\theta_i)}{1 - \Phi_i(\theta_i)}$  is the hazard rate of agent  $i$ . Substituting in the seller's payoff, we can write it as the expected "virtual surplus":

$$E_\theta \left[ \theta_0 y_0(\theta) + \sum_{i=1}^n \left( \theta_i - \frac{1}{h_i(\theta_i)} \right) y_i(\theta) \right].$$

Finally, for  $i \geq 1$ , let  $\nu_i(\theta_i) = \theta_i - 1/h_i(\theta_i)$ , which we will call the "virtual valuation" of agent  $i$ . Let also  $\nu_0(\theta_0) = \theta_0$ . Then the seller's program can be written as

$$\begin{aligned} \max_{x(\cdot)} E_\theta \quad & [\sum_{i=0}^n \nu_i(\theta_i) y_i(\theta_i)] \quad s.t. \quad \sum_{i=0}^n y_i(\theta) = 1, \\ & E_{\theta_{-i}} y_i(\theta_i, \theta_{-i}) \text{ is nondecreasing in } \theta_i, \forall i \geq 1. \quad (BM) \end{aligned}$$



If we ignore Bayesian Monotonicity (BM) above, we can maximize the above expectation for each state  $\theta$  independently. The solution then is to give the object to the agent who has the highest virtual valuation in the particular state: we have  $y_i(\theta) = 1$  when  $\nu_i(\theta_i) > \nu_j(\theta_j)$  for all agents  $j \neq i$ . The decision rule achieving this can be written as  $y(\theta) = y^*(\nu_0(\theta_0), \nu_1(\theta_1), \dots, \nu_n(\theta_n))$ , where  $y^*(\cdot)$  is the first-best efficient decision rule. Intuitively, the principal uses agents' virtual valuation rather than their true valuations because she cannot extract all of the agents' rents. An agent's virtual valuation is lower than his true valuation because it accounts for the agent's information rents which cannot be extracted by the principal. The profit-maximizing mechanism allocates the object to the agent with the highest virtual valuation.

Under what conditions can we ignore the monotonicity constraint? Note that when the virtual valuation function  $\nu_i(\theta_i) = \theta_i - 1/h_i(\theta_i)$  is an increasing function<sup>4</sup>, then an increase in an agent's valuation makes him more likely to receive the object in the solution to the relaxed problem. Therefore,  $y_i(\theta_i, \theta_{-i})$  is nondecreasing in  $\theta_i$  for all  $\theta_{-i}$ , and so by Proposition 15.10.2 the optimal allocation rule is implementable not just in BNE, but also in DS. The DIC transfers could be constructed by integration using DICFOC; the resulting transfers take a simple form:

$$t_i(\theta) = p_i(\theta_{-i})y_i(\theta), \text{ where } p_i(\theta_{-i}) = \inf\{\hat{\theta}_i \in [\underline{\theta}_i, \bar{\theta}_i] : y_i(\hat{\theta}_i, \theta_{-i}) = 1\}.$$

Thus, for each agent  $i$  there is a “pricing rule”  $p_i(\theta_{-i})$  that is a function of others' announcements, so that the agent wins if his announced valuation exceeds the price, and he pays the price whenever he wins. This makes truth-telling a dominant strategy by the same logic as in the Vickrey auction: lying cannot affect your own price but only whether you win or lose, and you want to win exactly when your valuation is above the price you face.

On the other hand, if virtual valuation functions are not increasing, (BM) may bind, and an “ironing” procedure should be used to identify the regions on which it binds.

Since the principal discounts agents' true valuations, she is less likely to sell the object than would be socially optimal. When she does sell the object, will it go to the agent with the highest valuation? This depends:

---

<sup>4</sup>A sufficient condition for this is that the hazard rate  $h_i(\theta_i)$  is a nondecreasing function. This is the same argument as in the one-agent case (note that  $\nu_{y\theta\theta} = 0$  for the linear utility function  $\nu(y, \theta) = \theta y$ ).

- The Symmetric Case: Agents' types are identically distributed. Then they all have the same virtual valuation function  $\nu(\cdot)$ . Suppose it is increasing. Then whenever the principal sells the good, it must go to the agent with the highest valuation  $\theta_i$ , because he also has the highest virtual valuation  $\nu(\theta_i)$ . The principal will sell the good when  $\nu(\max_{i \geq 1} \theta_i) > \theta_0$ .

The DS implementation of the optimal allocation rule is the Vickrey auction in which the seller submits a bid equal to  $p = \nu^{-1}(\theta_0)$ , called the "reserve price". It exceeds the principal's value  $\theta_0$  for the object, thus the object is less likely to be sold than in the social optimum. The intuition is that the principal reduces agents' consumption to reduce their information rents, just as in the one-agent case.

- The Asymmetric Case: The agent with the highest virtual valuation is now not necessarily the one with the highest true valuation. For example, suppose that we have two bidders, and  $h_1(\theta) < h_2(\theta)$  for all  $\theta$ . What this inequality means intuitively is that in an ascending price auction, where an agent drops out when the price reaches his true valuation, bidder 1 will be less likely to drop out at any moment conditional on both bidders staying in the game. Thus, the inequality means that bidder 1 is in a certain stochastic sense "more eager" to buy than bidder 2.<sup>5</sup> Then  $\nu_1(\theta) < \nu_2(\theta)$  for all  $\theta$ , i.e., the optimal auction should be biased in favor of the "less eager" agent 2. One intuition for this is that the bias increases "competition" among the bidders.

### 15.10.5 Correlated Types

Using the ideas of Cremer and McLean (1988), we can see that with correlated types, "almost anything is implementable". For simplicity consider the quasi-linear environment with two agents and  $\Theta_i = \{L, H\}$  for each  $i$  (Cremer-McLean have arbitrarily finite numbers of agents and types). Let the joint probability matrix of  $(\theta_1, \theta_2)$  be as follows:

$$\begin{array}{cc} \pi(L, L) & \pi(L, H) \\ \pi(H, L) & \pi(H, H) \end{array}$$

---

<sup>5</sup>In fact, it obtains whenever the distribution of bidder 1's valuation dominates that of bidder 2's valuation in the Monotone Likelihood Ratio ordering.

(Note that the entries must add up to 1.) Correlation means that the determinant of the matrix is nonzero. Suppose for definiteness that it is positive:

$$\pi(L, L)\pi(H, H) - \pi(L, H)\pi(H, L) > 0,$$

which corresponds to positive correlation (otherwise we could relabel  $L$  and  $H$  for one of the agents).

Also denote the marginal and conditional probabilities as follows:

$$\begin{aligned}\pi_i(\theta_i) &= \pi(\theta_i, L) + \pi(\theta_i, H), \\ \pi_{-i}(\theta_{-i} \mid \theta_i) &= \pi(\theta_i, \theta_{-i}) / \pi_i(\theta_i).\end{aligned}$$

Assume that the marginal densities  $\pi_i(\theta_i) > 0$  (otherwise, agent  $i$  has only one possible type).

Also, let  $\theta'$  denote the type other than  $\theta$ .

**Proposition 15.10.4** *For an arbitrary choice rule  $(y(\cdot), t_1(\cdot), t_2(\cdot))$ , there exists a BIC choice rule  $(y(\cdot), \tilde{t}_1(\cdot), \tilde{t}_2(\cdot))$  with the same decision rule  $y(\cdot)$  and the same interim expected transfers:*

$$E_{\theta_{-i}}[\tilde{t}_i(\theta_i, \theta_{-i}) \mid \theta_i] = E_{\theta_{-i}}[t_i(\theta_i, \theta_{-i}) \mid \theta_i], i = 1, 2, \theta_i = L, H.$$

Proof. Construct the new transfers as follows:

$$\tilde{t}_i(\theta_i, \theta_{-i}) = \begin{cases} t_i(\theta_i, \theta_{-i}) + \Delta \cdot \pi_{-i}(\theta'_i \mid \theta_i) & \text{when } \theta_{-i} = \theta_i, \\ t_i(\theta_i, \theta_{-i}) - \Delta \cdot \pi_{-i}(\theta_i \mid \theta_i) & \text{when } \theta_{-i} \neq \theta_i. \end{cases}$$

Then we have

$$E_{\theta_{-i}}[\tilde{t}_i(\theta_i, \theta_{-i}) \mid \theta_i] - E_{\theta_{-i}}[t_i(\theta_i, \theta_{-i}) \mid \theta_i] = \pi_{-i}(\theta_i \mid \theta_i)[\pi_{-i}(\theta'_i \mid \theta_i)\Delta] + \pi_{-i}(\theta'_i \mid \theta_i)[- \Delta \pi_i(\theta_i \mid \theta_i)].$$

At the same time, the new mechanism is BIC when  $\Delta > 0$  is large enough. Indeed, the equilibrium interim expected utilities are not affected, while the interim expected utility from lying and announcing  $\theta'_i \neq \theta_i$  is

$$\begin{aligned}\tilde{\Phi}_i(\theta' \mid \theta) &= \Phi_i(\theta' \mid \theta) + \pi_{-i}(\theta' \mid \theta)[\pi_{-i}(\theta \mid \theta')\Delta] + \pi_{-i}(\theta \mid \theta)[- \pi_{-i}(\theta' \mid \theta')\Delta] \\ &= \Phi_i(\theta' \mid \theta) + \frac{\pi(\theta', \theta)\pi(\theta, \theta') - \pi(\theta, \theta)\pi(\theta', \theta')}{\pi_i(\theta)\pi_i(\theta')}\Delta.\end{aligned}$$

By the correlation assumption, the fraction is negative, hence  $\tilde{\Phi}(\theta' \mid \theta)$  can be made as small as desired by choosing a large enough  $\Delta > 0$ .

Intuitively, the proof gives agent  $i$  a lottery contingent on the other type, whose expected value is zero if agent  $i$  is truthful but negative if he lies. If the stakes in this lottery are sufficiently high, BIC will be satisfied.

For example, suppose the principal wants to maximize profit. The first-best could be achieved by having the first-best decision rule  $y^*(\cdot)$  while eliminating the agents' information rents at the same time. Indeed, first choose transfers to be

$$t_i(\theta_i) = \hat{u}_i(\theta_i) - v_i(y^*(\theta_i, \theta_{-i}), \theta_i),$$

where  $\hat{u}_i(\theta_i)$  is the utility of agent  $i$  of type  $\theta_i$  if he withdraws from the mechanism. This makes all types' IRs bind (both interim and ex-post). Then we can construct the interim-equivalent BIC according to the above Proposition.

One problem with this argument is that when  $\Delta$  is large enough, lying will also constitute a BNE. But this problem can actually be avoided, because it turns out, that the principal can fully extract surplus even while preserving DIC.

**Proposition 15.10.5** *For any DIC choice rule choice rule  $(y(\cdot), t_1(\cdot), t_2(\cdot))$  there exists another DIC choice rule  $(y(\cdot), \tilde{t}_1(\cdot), \tilde{t}_2(\cdot))$  in which the interim IRs bind.*

Proof. Construct the new transfers as

$$\tilde{t}_i(\theta) = t_i(\theta) + h_i(\theta_{-i}),$$

To make agent  $i$ 's interim IR bind, we need

$$E_{\theta_{-i}}[U_i(\theta_i, \theta_{-i}) + h_i(\theta_{-i}) \mid \theta_i] = \hat{u}_i(\theta_i).$$

Writing this out for both types, we have

$$\begin{aligned} \pi_{-i}(L \mid L)h_i(L) + \pi_{-i}(H \mid L)h_i(H) &= \hat{u}_i(L) - E_{\theta_{-i}}[U_i(L, \theta_{-i}) \mid L], \\ \pi_{-i}(L \mid H)h_i(L) + \pi_{-i}(H \mid H)h_i(H) &= \hat{u}_i(H) - E_{\theta_{-i}}[U_i(H, \theta_{-i}) \mid H]. \end{aligned}$$

This is a system of linear equations, whose determinant is

$$\pi_{-i}(L \mid L)\pi_{-i}(H \mid H) - \pi_{-i}(H \mid L)\pi_{-i}(L \mid H) = \frac{\pi(L, L)\pi(H, H) - \pi(H, L)\pi(L, H)}{\pi_i(L)\pi_i(H)} > 0$$

by the correlation assumption, therefore a solution exists.

In particular, the principal can achieve the first-best by using a VCG mechanism and extracting all surplus from the agents.

Cremer and McLean formulate conditions for BNE and DS implementation for more than two agents and types. The condition for DS implementation is proved stronger than the one for BNE implementation, but both hold for generic joint distributions of types.

### 15.10.6 Ex Post Implementation

Standard auction theory assumes the seller knows the distribution  $F$  from which the buyers' valuations are drawn. For instance, consider a multi-unit auction when each buyer can buy at most 1 unit. The optimal auction then sells to buyers in order of their valuations while the virtual valuation  $\nu(\theta_i)$  exceeds the marginal cost  $c(X) - c(X - 1)$ . This can be represented by a supply curve  $\nu^{-1}(c(X) - c(X - 1))$ . In the linear-cost case, when the marginal cost is a constant  $c$ , there is no reason to use an auction and can post a price  $\nu^{-1}(c)$ . (The same is true if buyers are small and aggregate demand and hence equilibrium marginal cost are predictable.)

But if the distribution  $F$  is not known, the monopoly does not have a demand curve. I.e., suppose that buyers' valuations are i.i.d draws from an unknown distribution  $F$  over which the designer has a Bayesian prior. This makes buyers' types correlated. What mechanism should be used? We now impose ex post mechanism design. The optimal mechanism maximizes expected virtual surplus s.t. the relevant monotonicity constraint:

$$\begin{aligned} \max_{x(\cdot)} \quad & E_{\theta} [\sum_{i=1}^n \nu(\theta_i | \theta_{-i}) y_i(\theta) - c(\sum_{i=1}^n y_i(\theta))] \\ \text{s.t.} \quad & y_i(\theta_i, \theta_{-i}) \text{ is nondecreasing in } \theta_i, \forall i \geq 1. \quad (DM) \end{aligned}$$

Here  $\nu(\theta_i | \theta_{-i}) = \theta_i - \frac{f(\theta_i | \theta_{-i})}{1 - F(\theta_i | \theta_{-i})}$  is buyer  $i$ 's conditional virtual valuation/marginal revenue function (by symmetry it is the same for all buyers).

For simplicity focus on the case of constant marginal cost  $c$ . Then if we ignore (DM) the optimal mechanism sells to buyer  $i$  iff  $\nu(\theta_i | \theta_{-i}) \geq c$ . If the virtual utility  $\nu(\theta_i | \theta_{-i})$  is nondecreasing in  $\theta_i$ , then the mechanism will in fact satisfy (DM) and can be implemented by offering each buyer the optimal monopoly price  $p(\theta_{-i})$  given information gleaned from others' reports and selling iff his announced valuation is above the price. (Note that

agents make announcements first and then prices are determined.) As the number of agents  $n \rightarrow \infty$ ,  $p(\theta_{-i})$  will converge to the optimal monopoly price by the law of large numbers, and so the seller's profit will converge to that he would get if he knew  $F$ .

Note that the mechanism is fundamentally different from traditional auctions: here a bid of buyer  $i$  affects allocation to other bidders directly rather than only through buyer  $i$ 's allocation. This effect through the "informational linkage" among buyers, rather than the "technological linkage" in the standard auction through the seller's cost function or capacity constraint.

### 15.10.7 Ex post Implementation with Common Values

In the quasi-linear environment with common values, agent  $i$ 's payoff is  $v_i(x, \theta) + t_i$ , which depends directly on  $\theta_{-i}$ .

The problem of correlation comes up again. We can have a setup of independent common values, such as the "wallet game" where all agents care about  $\sum_i \theta_i$ . But the more natural model is the "mineral rights" model, in which types are i.i.d. signals conditional on the unknown true state, and thus correlated a priori. Most of the literature on common-value auctions is "positive" - solves for an equilibrium and compares several practical mechanisms, rather than looking for an optimal mechanism. This research was spurred by Milgrom and Weber. Recently, there has been new work on optimal design restricting attention to "ex post implementation", which imposes the following "ex post IC":

$$v_i(x(\theta), \theta) + t_i(\theta_i, \theta_{-i}) \geq v_i(x(\hat{\theta}_i, \theta_{-i}), \theta) + t_i(\hat{\theta}_i, \theta_{-i}), \forall \theta_i, \hat{\theta}_i \in \Theta_i, \forall \theta_{-i} \in \Theta_{-i}.$$

This means that truth-telling is optimal for agent  $i$  regardless of what types other agents have, but assuming they follow their equilibrium strategy, i.e., report truthfully. Without the assumption, the other agents' reports could be  $\theta'_{-i}$  while their true types are  $\theta_{-i}$ , in which case agent  $i$ 's utility would be  $v_i(x(\hat{\theta}_i, \theta'_{-i}), \theta_i, \theta_{-i}) + t_i(\hat{\theta}_i, \theta_{-i})$ , and it would be impossible to ensure that truth-telling is always optimal for agent  $i$ . Note that the assumption is not needed with private values, in which case ex post IC coincide with DIC. In addition, ex post implementation imposes ex post IR. See Dasgupta and Maskin (QJE 2000).

## Reference

- Abreu, R. and A. Sen (1990), "Subgame Perfect Implementation: A Necessary and Almost Sufficient Condition," *Journal of Economic Theory* 50, 285-299.
- Abreu, R. and A. Sen (1991), "Virtual Implementation in Nash Equilibrium," *Econometrica* 59, 997-1021.
- Bolton, P. and M. Dewatripont (2005), *Contract Theory*, MIT Press.
- Calsarniglia, X. (1977), "Decentralized Resource Allocation and Increasing Returns," *Journal of Economic Theory* 14, 263-283.
- Corchón, L. C., *Theory of Implementation of Socially Optimal Decisions in Economics*, Macmilian Press, Ltd., 1996.
- Danilov, V. (1992), "Implementation via Nash Equilibrium," *Econometrica*, 60, 43-56.
- Dasgupta, P, P. Hammond and E. Maskin (1979), "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility," *Review of Economic Studies* 46, 185-216.
- Duggan, J. (1993), "Virtual Implementation in Bayesian Equilibrium with Infinite Types," *Parts I and II*, Mimeo. rsity of Minnesota Press), 1987.
- Groves, T. and M. Loeb, Incentives and public inputs, *Journal of Public Economics* 4 (1975), 211-226.
- Groves, T. and Ledyard, J. (1977), "Optimal Allocation of Public Goods; A Solution to the Free Rider Problem," *Econometrica* 45(4), 783-811.
- Groves, T and Ledyard, J. (1985), "Incentive Compatibility Since 1972," *Chapter 2: Information, Incentive, and Economic Mechanisms*, T. Groves, R. Radner, and S. Reiter, eds., (Unive
- Hayek, F. A. von (1935), "The Present State of the Debate," in *Collectivist Economic Planning*, F. A. von Hayek, ed., (London), 201-243.

- Hayek, F. A. von (1945), "The Use of Knowledge in Society," *American Economic Review* 35, 519-530.
- Hurwicz, L. (1972), "On Informational Decentralized Systems," in *Decision and Organization*, Radner, R. and C. B. McGuire, eds., in Honor of J. Marschak, (North-Holland), 297-336.
- Hurwicz, L. (1979a), "Outcome Function Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Point," *Review of Economic Studies* XLVL (2), 397-419.
- Hurwicz, L. (1979b), "On Allocations Attainable Through Nash Equilibria," *Journal of Economic Theory* 21(1), 140-165.
- Hurwicz, L. (1979c), "Balanced Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Points for Two or More Agents," in *General Equilibrium, Growth, and Trade*, Jerry R. Green and Jose A. Scheinkman, eds., (Academic Press, New York).
- Hurwicz, L. (1979d), "Socialism and Incentives: Developing a Framework," *Journal of Comparative Economics* 3, 207-216.
- Hurwicz, L. 1986a), "Incentive Aspects of Decentralization," in *Handbook of Mathematical Economics*, K. J. Arrow and M. D. Intriligator, eds., Vol. III, (NorthHolland).
- Hurwicz, L. (1986b), "On Informational Decentralization and Efficiency in Resource Allocation Mechanism," in *Studies in Mathematical Economics*, S. Reiter, ed., Mathematical Association of America.
- Hurwicz, L. (1986c), "On the Implementation of Social Choice Rules in Irrational Societies," in *Social Choice and Public Decision Making Essays in Honor of Kenneth J. Arrow*, ed., Vol. I, (Cambridge University Press).
- Hurwicz, L. (1999), "Revisiting Externalities," *Journal of Public Economic Theory*, 1, 225-246.
- Hurwicz, L., Maskin, E. and Postlewaite, A. (1984), "Feasible Implementation of Social Choice Correspondences by Nash Equilibria," *Mimeo*.



- Hurwicz, L. and M. Walker (1990), "On the Generic Nonoptimality of Dominant-Strategy Allocation Mechanism: A General Theorem that Includes Pure Exchange Economies," *Econometrica* 58, 683-704.
- Jackson, M..O. (1991), "Bayesian Implementation," *Econometrica* 59, 461-477.
- Jordan, J. S. (1982), "The Competitive Allocation Process in Informationally Efficient Uniquely," *Journal of Economic Theory*, 28, 1-18.
- Kreps, D. (1994), "Signalling," in *Handbook of Game Theory* , R. J. Aumann and S. Hart, eds., Vol. II, (North-Holland).
- Laffont, J.J. and E. Maskin (1979), "A Differential Approach to Expected Maximizing Mechanisms," in *Aggregation and Revelation of Preferences*, J.J. Laffont, ed., (North Holland).
- Laffont, J. J. and D. Martimort, *The Theory of Incentives: The Principal-Agent Model*, (Princeton University Press), 2002.
- Lange, O. (1936-37), "On the Economic Theory of Socialism," *Review of Economic Studies* 4.
- Lange, O. (1938), *On the Economic Theory of Socialism*, (Philadelphia: Lippincott).
- Lange, O. (1942), "The Foundations of Welfare Economics," *Econometrica* 10, 215-228.
- Lange, O. and F. M. Taylor (1938), *On the Economic Theory of Socialism*, B.E. Lippincott, ed., (New York).
- Lerner, A. P. (1944), *Economics of Control*, (New York).
- Li, Q., S. Nakmura and G. Tian, (1995), "Nash Implementation of the Lindahl Correspondence with Decreasing Returns to Scale Technology," *International Economic Review* 36, 34-50.
- Liu, L., and G. Tian (1999), "A Characterization of Optimal Dominant Strategy Mechanisms," *Review of Economic Design* 4, 205-218.

- Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic*, Oxford University Press, 1995, Chapter 14.
- E. Maskin (1999), "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies* 66, 23-38.
- Matsushima, H. (1988), "A New Approach to the Implementation Problem," *Journal of Economic Theory* 45, 128-144.
- Matsushima, H. (1993), "Bayesian Monotonicity with Side Payments," *Journal of Economic Theory* 59, 107-121.
- McKelvey, R. (1989), "Game Forms for Nash Implementation of General Social Correspondence," *Social Choice and Welfare*, 6, 139-156.
- Meng, D. and G. Tian, 'Multi-task Incentive Contract and Performance Measurement with Multidimensional Types,' *Games and Economic Behavior*, 77 (2013) 377-404.
- Mookherjee, D. and S. Reichelstein (1990), "Implementation via Augmented Revelation Mechanisms," *Review of Economic Studies* 57, 453-475.
- Moore, J. and R. Repullo (1988), "Subgame Perfect Implementation," *Econometrica* 56, 1191-1220.
- Moore, J. and R. Repullo (1990), "Nash Implementation: A Full Characterization," *Econometrica* 56, 1083-1099.
- Mount, K., and S. Reiter (1974), "Informational Size of Message Spaces," *Journal of Economic Theory* 8, 161-191.
- Palfrey, T. and S. Srivastava (1987), "On Bayesian Implementable Allocations," *Review of Economic Studies* 54, 193-208.
- Palfrey, T. and S. Srivastava (1989), "Mechanism Design with Incomplete Information: A Solution to the Implementation Problem," *Journal of Political Economy* 97, 668-691.
- Palfrey, T. and S. Srivastava (1989), "Implementation with Incomplete Information in Exchange Economies," *Econometrica* 57, 115-134.

- Palfrey, T. and S. Srivastava (1991), "Nash Implementation Using Undominated Strategy," *Econometrica* 59, 479-502.
- Postlewaite, A. (1985), Implementation in Nash Equilibria in Economic Environments," in *Social Goal and Social Organization Essays in Memory of Elisha Pazner*, L. Hurwicz, D. Schmeidler, and H. Sonnenschein, eds.,(Cambridge University Press).
- Postlewaite, A. and D. Schmeidler (1986), "Implementation in Differential Information Economies," *Journal of Economic Theory* 39, 14-33.
- Postlewaite, A. and D. Wettstein (1989), "Continuous and Feasible Implementation," *Review of Economic Studies* 56, 603-611.
- Repullo, R.(1986), "The Revelation Principle under Complete and Incomplete Information," in *Economic Organization as Games*, Binmore, K. and P. Dasgupta, eds., (Oxford, Basil Blackwell).
- Repullo, R.(1987), "A Simple Proof of Maskin's Theorem on Nash Implementation," *Social Choice and Welfare*, 4, 39-42.
- Schmeidler, D. (1980), "Walrasian Analysis via Strategic Outcome Functions," *Econometrica* 48, 1585-1593.
- Segal, I. (2007), *Lecture Notes in Contract Theory*, Stanford University.
- Tian, G. (1989), "Implementation of the Lindahl Correspondence by a Single-Valued, Feasible, and Continuous Mechanism," *Review of Economic Studies* 56, 613-621.
- Tian, G. (1990), "Completely Feasible and Continuous Nash-Implementation of the Lindahl Correspondence with a Message Space of Minimal Dimension," *Journal of Economic Theory* 51, 443-452.
- Tian, G. (1991), "Implementation of Lindahl Allocations with Nontotal-Nontransitive Preferences," *Journal of Public Economics* 46, 247-259.
- Tian, G. (1992), "Implementation of the Walrasian Correspondence without Continuous, Convex, and Ordered Preferences," *Social Choice and Welfare* 9, 117- 130.

- Tian, G. (1993), "Implementing Lindahl Allocations by a Withholding Mechanism," *Journal of Mathematical Economics* 22, 169-179.
- Tian, G. (1994), "Implementation of Linear Cost Share Equilibrium Allocations," *Journal of Economic Theory* 64, 568-584.
- Tian G. (1994a), "On Informational Efficiency and Incentive Aspects of Generalized Ratio Equilibria," *Journal of Mathematical Economics* 23, 323-337.
- Tian, G. (1994b), "Implementation of Linear Cost Share Equilibrium Allocations," *Journal of Economic Theory* 64, 568-584.
- Tian, G. (1996a), "On the Existence of Optimal Truth-Dominant Mechanisms," *Economics Letters* 53, 17-24.
- Tian, G. (1996b), "Continuous and Feasible Implementation of Rational Expectation Lindahl Allocations," *Games and Economic Behavior* 16, 135-151.
- Tian, G. (1997), "Virtual Implementation in Incomplete Information Environments with General Sets of Alternatives and Types," *Journal of Mathematical Economics* 28, 313-339.
- Tian, G. (1999a), "Double Implementation in Economies with Production Technologies Unknown to the Designer," *Economic Theory*, 13 (1999), 689-707.
- Tian, G. (1999b), "Bayesian Implementation in Exchange Economies with State Dependent Preferences and Feasible Sets," *Social Choice and Welfare*, 16 (1999), 99-119.
- Tian, G. (2000a), "Double Implementation of Lindahl Allocations by a Continuous and Feasible Mechanism," *Social Choice and Welfare*, 17 (2000), 125-141.
- Tian, G. (2000b), "Incentive Mechanism Design for Production Economies with Both Private and Public Ownership," *Games and Economic Behavior*, 33 (2000), 294-320.
- Tian, G. (2000c), "Implementation of Balanced Linear Cost Share Equilibrium Solution in Nash and Strong Nash Equilibria," *Journal of Public Economics*, 76 (2000), 239-261.

- Tian, G. (2000d), "Double Implementation of Linear Cost Share Equilibrium Allocations," *Mathematical Social Sciences*, 40 (2000), 175-189.
- Tian, G. (2000e), "A Unique Informationally Efficient Allocation Mechanism in Economies with Public Goods," *Mimeo*.
- Tian, G. (2004), "A Unique Informationally Efficient Allocation Mechanism in Economies with Consumption Externalities," *International Economic Review*, 45 (2004), 79-111.
- Tian, G. (2005), "Implementation in Production Economies with Increasing Returns," *Mathematical Social Sciences*, forthcoming.
- Tian, G. (2006), "The Unique Informational Efficiency of the Competitive Mechanism In Economies with Production," *Social Choice and Welfare*, 26 (2006), 155-182.
- Tian, G. (2009a), "Implementation in Economies with Non-Convex Production Technologies Unknown to the Designer" *Games and Economic Behavior*, 66 (2009), 526-545.
- Tian, G. (2009b), "Implementation of Pareto Efficient Allocations," *Journal of Mathematical Economics*, 45 (2009), 113-123.
- Tian, G. (2009c) "Implementation of Marginal Cost Pricing Equilibrium Allocations with Transfers in Economies with Increasing Returns to Scale," *Review of Economic Design*, forthcoming.
- Tian, G. and Q. Li (1991), "Completely Feasible and Continuous Implementation of the Lindahl Correspondence with Any Number of Goods," *Mathematical Social Sciences* 21, 67-79.
- Tian, G. and Q. Li (1994), "An Implementable and Informational Efficient State-Ownership System with General Variable Returns," *Journal of Economic Theory* 64, 268-297.
- Tian, G. and Q. Li (1995a), "On Nash-Implementation in the Presence of Withholding," *Games and Economic Behavior* 9, 222-233.

- Tian, G. and Q. Li (1995b), "Ratio-Lindahl Equilibria and an Informationally Efficient and Implementable Mixed-Ownership System," *Journal of Economic Behavior and Organization* 26, 391-411.
- Thomson, W. (1979) "Maximum Strategies and Elicitation of Preferences," in *Aggregation and Revelation of Preferences*, J. J. Laffont, ed., (North-Holland).
- Varian, H.R. (1992), *Microeconomic Analysis*, (W.W. Norton and Company, Third Edition).
- Walker, M. (1978), "A Note on the Characterization of Mechanisms for the Revelation of Preferences," *Econometrica* 46, 147-152.
- Walker, M. (1980), "On the Nonexistence of a Dominant Strategy Mechanism for Making Optimal Public Decisions," *Econometrica* 48, 1521-1540.
- Walker, M. (1981), "A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations," *Econometrica* 49, 65-71.
- Wolfstetter, E., *Topics in Microeconomics - Industrial Organization, Auctions, and Incentives*, Cambridge Press, 1999, Chapters 8-10.

# Chapter 16

## Dynamic Mechanism Design

Chapters 13 – 15 deal only with static or one-shot contracting situations. In practice, many relations are repeated or long term. In this chapter we will study dynamic contract theory.

A number of new fundamental economic issues arise when the parties are involved in a long-term contractual relation. How is private information revealed over time? How is the constrained efficiency of contractual outcomes affected by repeated interactions? How does the possibility of renegotiation limit the efficiency of the overall long-term contract? To what extent can *reputation* serve as a more informal enforcement vehicle that is an alternative to courts? We discuss some of these issues in this chapter.

We will study long-term incentive contracting in a dynamic principal-agent setting with one-agent and adverse selection. We first assume that the principal (designer) can commit to a contract forever, and then consider what happens when she cannot commit against modifying the contract as new information arrives.

### 16.1 Dynamic Contracts with Commitment

Methodologically, there is no significant change in analyzing optimal multiperiod contracts as long as the contracting parties can *commit* to a single comprehensive long-term contract at the initial negotiation stage. When full commitment is feasible, the long-term contract can essentially be reduced to a slightly more complex static contract involving trade of a slightly richer basket of state-contingent commodities, services, and transfers. What

this conclusion implies in particular for contracting under hidden information is that the revelation principle still applies under full commitment.

### 16.1.1 Dynamic P-A Model with Constant Type

For this purpose, we consider a principal-agent relationship that is repeated over time, over a finite or infinite horizon. First, as a benchmark, we suppose that the agent's type  $\theta$  is realized ex ante and is then constant over time.

An outcome in this model can be described as a sequence of decisions and payments  $(x_t, y_t)_{t=0}^{\infty}$ . Both parties have the same discount factor  $0 < \delta < 1$  (in a finite-horizon model, discounting does not matter). Suppose the payoffs are stationary:

$$\begin{aligned} \text{Principal: } & \sum_{t=0}^{\infty} \delta^t [y_t - c(x_t)], \\ \text{Agent: } & \sum_{t=0}^{\infty} \delta^t [v(x_t, \theta) - y_t]. \end{aligned}$$

Suppose the Principal can commit ex ante to a mechanism, which cannot be modified later on. This can be understood as a standard mechanism design setting in which the outcome is the whole sequence of decisions and payments. By the Revelation Principle, we can restrict attention to direct revelation mechanisms, in which the agent announces his type  $\theta$ , and the mechanism implements a sequence  $(x_t(\theta), y_t(\theta))_{t=0}^{\infty}$  according to his announcement.

Does the principal need to use mechanisms that are not stationary, i.e., specify different outcomes in the two periods? The answer is no: An outcome can be regarded as a stationary randomization (lottery) that yields, in each period each  $(x_t, t_t)$  with probability  $\frac{\delta^t}{1-\delta}$ . Then the parties' expected utilities will be the same as from the original non-stationary outcome. Thus, the stationary setting with constant types and commitment of the optimal one-period randomizes contract.

Under some conditions, the principal will not resort to lotteries. A simple case is where the functions  $c(\cdot)$  and  $v(\cdot, \theta)$  are linear in  $x$ : In this case the randomized outcome  $(\tilde{x}, \tilde{t})$  can be replaced with a deterministic outcome  $(E\tilde{x}, E\tilde{t})$ , which gives both parties exactly the same utilities. More generally, a set of sufficient conditions for the principal not resort to randomizations is given by  $c''(\cdot) \geq 0$ ,  $v_{xx}(\cdot) \leq 0$ , and  $v_{\theta xx}(\cdot) \geq 0$  (see Fudenberg-Tirole pp. 306-7).



### 16.1.2 Dynamic P-A Model with Markov Types

Suppose now type can change: let  $\theta_t$  denote the type in period  $t$ , and suppose that types follow a Markov process. This captures many dynamic contracting situations in which agents learn information over time such as:

- Consumers buy calling plans and then learn their valuations for calls on a given day.
- Travelers buy tickets and then get new information about their valuation for the trip (e.g., sickness, busy schedule, etc.)
- A buyer joins a wholesale club for a fee and then learns his valuations for goods.

$$\begin{aligned}\text{Principal: } & \sum_{t=0}^{\infty} \delta^t [y_t - c(x_t)], \\ \text{Agent: } & \sum_{t=0}^{\infty} \delta^t [v(x_t, \theta_t) - y_t].\end{aligned}$$

We now allow for non-stationary payoffs now (and in particular finite horizon, or even perhaps trade does not occur until some period  $T$  even though information arrives earlier and the agents sends messages.

The model has “multidimensional type” and (even though with restriction that an agent cannot report a type before he observes it), and so it is quite complicated.

How do optimal screening mechanisms look in such situations? We can still use a version of the Revelation Principle for dynamic games to restrict attention to mechanisms in which the Agent reports new information as it arrives and reports truthfully in equilibrium. The mechanism then specifies, for each  $t$ , each  $x_t(\theta^t)$  and  $y_t(\theta^t)$ , where  $\theta^t = (\theta_1, \dots, \theta_t)$  is the reporting history at time  $t$ .

**Remark 16.1.1** The agent’s complete strategy specifies what to report if he lied in the past, but the Markovian assumption means that his true past is no longer payoff-relevant (even if his past reports may be), so if he finds it optimal to be report truthfully after truthful reporting he will also find it optimal to report truthfully after misreporting. Using the “one-stage deviation principle”, we will check the IC that the agent does not want to deviate to lying about  $\theta_t$  for any history  $\theta^{t-1}$ .

Let  $U_t(\theta^t)$  denote continuation expected utility of type  $\theta_t$  in period  $t$  from truthful announcement for reporting history  $\theta^{t-1}$ .

Consider first the IRs. It is easy to see that even if we require participation constraints in all periods (i.e.,  $U_t(\theta^t) \geq 0, \forall t, \theta^t$ ), only zero-period IRs can bind: Any other IR in period  $t > 1$  could be relaxed by “bonding” the agent - i.e., charging him a large payment  $M$  in period zero and returning it in period  $t + 1$ . (There are real-life examples of such bonding: e.g., long-term life insurance contracts, which consumers can cancel at any time, have premiums higher than risk in the beginning of the contract, in exchange for premia below risk later in life. This ensures that consumers who find out that they are healthy and likely to live long don’t cancel the contract.)

Since only 1st-period IR may bind, what determines distortion is extraction by the principal of the agent’s zero-period info rent  $U_0(\theta_0)$ .

To calculate this rent, suppose that all  $\theta_t \in [\underline{\theta}, \bar{\theta}]$ , the Markov process is described by a smooth cdf  $F_t(\theta_t | \theta_{t-1})$ , and  $v_t$  has a bounded derivative in  $\theta_t$ . Then we can get ICs for 1-stage deviations from truth-telling as follows:

$$U_t(\theta^t) = \max_{\hat{\theta}_t} \left[ v_t(x_t(\theta^{t-1}, \hat{\theta}_t), \theta_t) + \delta \int U_{t+1}(\theta_{t-1}, \hat{\theta}_t, \tilde{\theta}_{t+1}) dF_{t+1}(\tilde{\theta}_{t+1} | \theta_t) \right].$$

Applying the Envelope Theorem yields (a.e.):

$$\begin{aligned} \frac{\partial U_t(\theta^t)}{\partial \theta_t} &= \frac{\partial v_t(x_t(\theta^t), \theta_t)}{\partial \theta_t} + \delta \int U_{t+1}(\theta^{t+1}) \frac{\partial f_{t+1}(\theta_{t+1} | \theta_t)}{\partial \theta_t} d\theta_{t+1} \\ &= \frac{\partial v_t(x_t(\theta^t), \theta_t)}{\partial \theta_t} - \delta \int U_{t+1}(\theta^{t+1}) \frac{\partial F_{t+1}(\theta_{t+1} | \theta_t)}{\partial \theta_{t+1}} d\theta_{t+1} \\ &= \frac{\partial v_t(x_t(\theta^t), \theta_t)}{\partial \theta_t} - \delta E \left[ \frac{U_{t+1}(\theta^{t+1})}{\frac{\partial F_{t+1}(\theta_{t+1} | \theta_t)}{\partial \theta_{t+1}}} \mid \theta_t \right]. \end{aligned}$$

(where the 2nd equality obtains by integration by parts using the fact that  $\frac{\partial F_{t+1}(\theta_{t+1} | \theta_t)}{\partial \theta_{t+1}} = 0$  for  $\theta_{t+1} = \underline{\theta}$  and  $\theta_{t+1} = \bar{\theta}$ ).

So, working backward from the infinite horizon and applying iterated expectations, we get

$$U'_0(\theta_0) = E \sum_{t=0}^{\infty} \delta^t \Pi_{\tau=1}^t \left( - \frac{\partial F_t(\theta_t | \theta_{t-1}) / \partial \theta_{t-1}}{f_t(\theta_t | \theta_{t-1})} \right) \frac{\partial v_t(x_t(\theta^t), \theta_t)}{\partial \theta_t}.$$

We need to figure out which of the period-1 IRs binds. Assume FOSD:  $\partial F_t(\theta_t | \theta_{t-1}) / \partial \theta_{t-1} \leq 0$ . Then  $\partial U_0(\theta_0) / \partial \theta_0 \geq 0$ , and so only IR binds. Intuitively, the distortion

will then be in the direction of reducing  $U'_0(\theta_0)$ , since this reduces all types' information rents. Under FOSD, this is achieved by reducing the derivative  $\frac{\partial v_t(x_t(\theta^t), \theta_t)}{\partial \theta_t}$ , which under SCP is achieved by reducing  $x_t(\theta^t)$ . To see this more formally, we can write virtual surplus in period  $t$  in state  $\theta$  as

$$v_t(x_t(\theta^t), \theta_t) - c_t(x_t) - \frac{1 - F_0(\theta_0)}{f(\theta)} \Pi_{\tau=1}^t \left( -\frac{\partial F_\tau(\theta_\tau | \theta_{\tau-1})}{\partial \theta_{\tau-1}} \right) \frac{\partial v_t(x_t(\theta^t), \theta_t)}{\partial \theta_t}.$$

The relaxed problem maximizes this expression in all periods and states. Properties of solution to the relaxed problem:

- Downward distortion in each period.
- For  $\theta_0 = \bar{\theta}$ , no distortions forever.
- For  $t \geq 1$ , if  $\theta_t = \underline{\theta}$  or  $\theta_t = \bar{\theta}$ , then no distortions from then on (since then  $\partial F_t(\theta_t | \theta_{t-1}) / \partial \theta_{t-1} = 0$ ).

Does this solve the full problem - are all ICs satisfied? In 2-period model with linear utility and FOSD, Courty-Li show that all ICs are satisfied when  $x_t(\theta^t)$  is nondecreasing in all arguments, and that this is satisfied when  $\theta_t = a^{-1}(b(\theta_{t-1}) + \varepsilon_t)$  (including e.g. additive and multiplicative case), then  $F(\theta_t | \theta_{t-1}) = G(a(\theta_t) - b(\theta_{t-1}))$ , and so  $F_{\theta_{t-1}}(\theta_t | \theta_{t-1}) / f(\theta_t | \theta_{t-1}) = -b'(\theta_{t-1}) / a'(\theta_t)$ . So when  $b(\cdot)$  is increasing and convex and  $a(\cdot)$  is increasing and concave we are OK.

Discussion about dynamic multi-agent model can be seen in Athey and Segal (2007).

## 16.2 Dynamic Contracts without Full Commitment

When the contracting parties are allowed to *renegotiate* the initial contract as time unfolds and new information arrives, new conceptual issues need to be addressed and the basic methodology of optimal static contracting must be adapted. Mainly, incentive constraints must then be replaced by tighter *renegotiation-proofness* constraints.

### 16.2.1 Basic Problem

Now we examine the temptations to modify the original contract and how they affect the outcome with imperfect commitment. For simplicity we focus on a 2-period P-A model

with a constant type and stationary payoffs. That is, the payoffs a

$$\text{Principal: } y_1 - c(x_1) + \delta[y_2 - c(x_2)],$$

$$\text{Agent: } v(x_1, \theta) - y_1 + \delta[v(x_2, \theta) - y_2].$$

The agent's type  $\theta$  is realized before the relationship starts and is the same in both periods. Suppose the type space is  $\Theta = \{L, H\}$ , and let  $\Pr(\theta = H) = \mu \in (0, 1)$ . We will allow  $\delta$  to be smaller or greater than one—the latter can be interpreted as capturing situations in which the second period is “longer” than the first period (we avoid analyzing more than 2 periods, since the model proves already complicated as it is).

The outcome of contracting in this model depends on the degree of the principal's ability to commit not to modify the contract after the first period.

If the principal cannot commit, however, she will be tempted to modify the optimal commitment contract using the information revealed by the agents in the first period. Recall that the optimal commitment contract is a repetition of the optimal static contract.

There are two reasons for the principal to modify the contract:

1. The high type receives information rents in the optimal contract (point H). However, after the first period, the high type has already revealed himself. Then the principal can extract all the information rents from the high type by offering point  $H'$  instead. This is called the *Ratchet Problem*. Intuitively, once the agent has revealed that he has a high value for consumption, the principal charges more. If the high type anticipates this modification ex ante, his ex ante incentive constraint becomes violated: he would rather pretend being the low type in the first period by choosing point  $L$ . This will not change his first-period payoff (since  $(IC_{HL})$  binds), but allows him to keep information rents in the 2nd period. Examples of ratchet problems: labor contracts, regulation of natural monopolies, etc.
2. After a low type reveals himself by choosing point  $L$ , the principal can raise her second-period payoff by eliminating the downward distortion in the low type's consumption in the second period. Namely, instead of point  $L$  she will offer the low type the efficient point  $L'$ . This is called the *renegotiation problem*, because this modification can be done with the agreement of the low type. (In fact, the low type buyer will be indifferent between  $L$  and  $L'$ , but the principal can make her strictly

prefer the change by asking for  $\varepsilon$  less in the transfer.) The renegotiation problem was first considered by Dewatripont (1988, 1989).

If the high type anticipates this modification, in the first period he will pretend being the low type by choosing point  $L$  instead of point  $H$ , and then receiving point  $L'$  in the second period. This will not change his first-period payoff (since  $(IC_{HL})$  binds), but will increase his second-period payoff, since  $L'$  gives him a higher payoff than  $H$ .

Examples of renegotiation problems: soft budget constraints.

Both these ex post modifications improve the principal's payoff post, however, they destroy the ex ante incentive-compatibility of the contract.

Observe that in an equilibrium without the principal's commitment, the principal cannot do better than if she could commit. Indeed, any choice rule that results in an equilibrium without commitment will satisfy the agent's IC and IR constraints, therefore it can also be implemented by an IC and IR commitment mechanism. The reverse is not true: the principal's inability to commit will limit the set of implementable choice rules and reduce her expected ex ante profits.

To what extent can the principal commit not to modify the contract and avoid the ratchet and renegotiation problems? The literature has considered three degrees of commitment:

1. **Full Commitment:** The principal can commit to any contract ex-ante. We have already considered this case. The Revelation Principle works, and we obtain a simple replication of the static model.
2. **Long Term Renegotiable Contracts:** The principal cannot modify the contract unilaterally, but the contract can be renegotiated if both the principal and the agent agree to it. Thus, the principal cannot make any modifications that make the agent worse off, but can make modifications that make the agent better off. Thus, the ratchet problem does not arise in this setting (the high type would not accept a modification making him worse off) but the renegotiation problem does arise.
3. **Short-Term Contracts:** The principal can modify the contract unilaterally after the first period. This means that the ex ante contract has no effect in the second

period- i.e., it can only specify the first-period outcome, and after the first period the parties contract on the second-period outcome. This setting gives rise to both the ratchet problem and the renegotiation problem.

A key feature of the cases without commitment is that when  $\delta$  is sufficiently high, the principal will no longer want the agent to reveal his type fully in the first period she prefers to commit herself against contract modification by having less information at the modification stage. For example, if the two types pool in the first period, the principal can implement the optimal static contract in the second period, and the problems of ratcheting and renegotiation do not arise. When  $\delta$  is very high, this will indeed be optimal for the principal. For intermediate levels of  $\delta$ , the principal will prefer to have partial revelation of information in the first period, which allows her to commit herself against modification. This partial revelation will be achieved by the agent using a mixed strategy, revealing his type with some probability and pooling with some probability.

### 16.2.2 Full Analysis of A Simple Model

Since the analysis of models without full commitment is quite difficult, we will study the simplest possible model in which the ratchet and renegotiation problems arise. In the model, based on Hart and Tirole (1990), the principal is a monopolistic seller who can produce an indivisible object in each period at a zero cost. The agent's valuation for this object in each period is  $\theta$ , thus his per-period utility is  $\theta x - t$ , where  $x \in \{0, 1\}$  is indicates whether he receives the object, and  $\theta \in \{L, H\}$  is his privately observed type, with  $0 < L < H$ , and  $\Pr(\theta = H) = \mu \in (0, 1)$ . This is a particular case of the two-type principal-agent model considered before.

#### The optimal static contract

The simplest way to solve for the optimal static contract is using the taxation principle. Suppose the principal offers a tariff  $T : \{0, 1\} \rightarrow \mathbb{R}$ . Since the agent can reject the tariff, we can assume that  $T(0) = 0$ . Thus, we can look for the optimal  $T(1) = p$  - price for the object.

Faced with the tariff, the agent solves  $\max_{x \in \{0, 1\}} (\theta x - px)$ . Thus, he chooses  $x(\theta) = 1$  if  $\theta > p$  and  $x(\theta) = 0$  if  $\theta < p$ . Let  $\bar{x} = \mu x(H) + (1 - \mu)x(L)$  - the expected demand faced

by the principal. Then (assuming the principal can convince the agent to buy when he is indifferent) we have

$$\bar{x}(p) = \begin{cases} 0 & \text{if } p > H, \\ \mu & \text{if } p \in (L, H], \\ 1 & \text{if } p \leq L. \end{cases}$$

The principal will choose  $p$  to maximize expected revenue  $p\bar{x}(p)$ . There are only two potentially optimal alternatives: setting  $p = L$  and selling to both types yields expected revenue  $L$ , and setting  $p = H$  and selling only to the high type yields expected revenue  $\mu H$ . Thus, the principal's optimal price is

$$p^*(\mu) = \begin{cases} L & \text{if } \mu < L/H, \\ H & \text{if } \mu > L/H, \\ \text{either } L \text{ or } H & \text{if } \mu = L/H. \end{cases}$$

Intuitively, if there are sufficiently many high types, the principal prefers to target only them by setting price  $H$  and making the sale with probability  $\mu$ , but if there are few high types the principal sells price  $L$  and sells with probability 1. (Exercise: obtain the same result using the Revelation Principle instead of the Taxation Principle.)

Note that when  $p^*(\mu) = H$ , both types consume efficiently, but the high type receives information rents. When  $p^*(\mu) = L$ , on the other hand, the low type underconsumes, but the high type receives no information rent. The slight difference from the basic principal-agent model is that the agent's consumption choices are never interior, so the high type's information rents and the low type's underconsumption need not always be strictly positive. However, the basic trade-off between efficiency of the low type's consumption and extraction of the high type's information rents is the same as in the basic model.

### Dynamics with Full Commitment

Suppose now that there are two periods. We already know that any two-period contract can be replaced with a repetition of the same (possibly randomized) static contract. In our case, the optimal static contract is setting the price  $p^*(\mu)$ . (Since the parties' payoffs are linear both  $x$  and  $t$ , randomizations need not be used.) Therefore, the optimal commitment contract sets the same price  $p^*(\mu) = L$  in both periods.

If the principal lacks full commitment, will she want to modify this contract after the first period? This depends on  $p^*(\mu)$ . If  $\mu < L/H$  and therefore  $p^*(\mu) = L$ , then the two types pool in the first period, and the principal receives no information. Then she has no incentive to modify the same optimal static contract in the second period.

However, when  $\mu > L/H$  and therefore  $p^*(\mu) = H$ , the principal knows following a purchase that she deals with a high type and following no purchase that she deals with a low type. In the latter case, she will want to restore efficiency for the low type by the price to  $L$  in the second period - the renegotiation problem. Anticipating this price reduction following no purchase, the high type will not buy in the first period. As for the former case, the principal already extracts all of the high type's information rents, so the ratcheting problem does not arise at the optimal commitment contract. We will see, however, that when the principal deals optimally with the renegotiation problem, the ratcheting problem may arise and further hurt the principal.

## Dynamics with Short-Term Contracts

Let  $\mu > L/H$ , and let the parties play the following two-period game:

- Principal offers price  $p_1$ .
- Agent chooses  $x_1 \in \{0, 1\}$ .
- Principal offers price  $p_2$ .
- Agent chooses  $x_2 \in \{0, 1\}$ .

We solve for the principal's preferred weak Perfect Bayesian Equilibrium (PBE) in this extensive-form game of incomplete information. We do this by first considering possible continuation PBEs following different first-period price choice  $p_1$ . Then the principal's optimal PB can be constructed by choosing the optimal continuation PBE following any price choice  $p_1$ , and finding the price  $p_1^*$  that gives rise to the best continuation PBE for the principal.

The principal's strategy in the continuation game is a function  $p_2(x_1)$  that sets the second-period price following the agent's first period choice  $x_1$ . This price should be optimal given  $\mu_2(x_1) \equiv \Pr(\theta = H \mid x_1)$  - the principal's posterior belief that the agent is



a high type after  $x_1$  was chosen. From the analysis of the static model we know that the principal's 2nd period price must satisfy

$$p_2(x_1) = p^*(\mu_2(x_1)) = \begin{cases} L & \text{if } \mu_2(x_1) < L/H, \\ H & \text{if } \mu_2(x_1) > L/H, \\ \text{either } L \text{ or } H & \text{if } \mu_2(x_1) = L/H. \end{cases}$$

**“Full Revelation Equilibrium”** Can we have a continuation PBE in which  $x_1(H) = 1$  and  $x_1(L) = 0$ ? In this case we have  $\mu_2(0) = 0$  and  $\mu_2(1) = 1$ , and therefore  $p_1(0) = L$  and  $p_2(1) = H$ . What should be the first-period price  $p_1$  for this to happen? If the high type does not buy in the first period, this reduces the price he pays in the second period. Thus, in order to induce him to buy, the first-period price must be sufficiently low. More precisely, we must have

$$\underbrace{H - p_1}_{\text{buy at } \tau = 1} + \underbrace{\delta(H - H)}_{\text{no rents at } \tau = 2} \geq \underbrace{0}_{\text{don't buy at } \tau = 1} + \underbrace{\delta(H - L)}_{\text{rents at } \tau = 2},$$

or

$$p_1 \leq (1 - \delta)H + \delta L.$$

As for the low type, he is willing to pay exactly  $L$  in the first period (he receives no surplus in the 2nd period in any event). Thus, when  $\delta \in (0, 1)$ , the optimal “Full Revelation” equilibrium has the principal selling to the high type only in the first period at price  $p_1 = (1 - \delta)H + \delta L \in (L, H)$ , and the low type does not buy. The principal expected profit is

$$\pi_F = \mu[(1 - \delta)H + \delta L + \delta H] + (1 - \mu)\delta L = \mu H + \delta L.$$

However, when  $\delta > 1$ , the principal cannot sell to the high type in the first period without selling to the low type, since  $(1 - \delta)H + \delta L < L$ . Then the Principal cannot achieve Full Revelation with short-term contracts. Intuitively, when the future is important, the ratchet effect is so large that the high type needs to be offered a good deal to reveal himself in the first period. But then the low type would “take the money and run”. This makes full revelation impossible.

**“No Revelation Equilibrium”** Another possibility for the principal is to have the agent always make the same choice in the first period. The principal's second-period

belief in equilibrium is  $\mu$ , and the second-period equilibrium price is therefore  $H$ . Note that the low type will be willing to pay price  $L$  in the first period (he receives no surplus in the second period in any event). Thus, given that there is no revelation, it will be optimal for the principal to sell to both agents in the first period at price  $L$ . Since  $x_1 = 0$  is a zero-probability event, the belief  $\mu_2(0)$  can be chosen arbitrarily. In particular, we can choose  $\mu_2(0) > L/H$ , so that  $p_2(0) = p_2(1) = H$ . The high type will then also be willing to buy at price  $L$  (he would not gain anything by not buying). The principal's expected profit in this candidate equilibrium is

$$\pi_N = L + \delta\mu H.$$

Which continuation equilibrium is better for the principal? Note that

$$\pi_F - \pi_N = (\mu H - L)(1 - \delta).$$

Recall that we now consider the case where  $\mu > L/H$ . In this case, when  $\delta < 1$ , Full Revelation is feasible and it is better for the principal than No Revelation. On the other hand, recall that when  $\delta > 1$ , Full Revelation is not feasible at all.

**"Partial Revelation Equilibrium"** The principal may be able to do better than either with Full Revelation or Partial Revelation by letting the agent use a mixed strategy. In such mixed-strategy equilibrium, the agent's first-period choice gives some information about the agent type but does not reveal it fully. Thus, let the agent randomize in the first period, with denoting the probability that the agent of type chooses  $x_1 = 1$ . We will consider equilibria in which both first- period choices are made with a positive probability, hence both  $\mu_2(0)$  and  $\mu_2(1)$  are obtained by Bayes' rule (so the "No Revelation" case is ruled out). There are four cases to consider:

1.  $p_2(0) = p_2(1) = L$
2.  $p_2(0) = H, p_2(1) = L$
3.  $p_2(0) = L, p_2(1) = H$
4.  $p_2(0) = p_2(1) = H$

We consider these four cases in turn.

**Case 1** In this case we must have  $\mu_2(0) \leq L/H$  and  $\mu_2(1) \leq L/H$ . This is inconsistent with the posteriors being obtained by Bayes rule from the prior  $\mu > L/H$ .

**Case 2** This case is also infeasible, though the reasoning is more complex. Note that the payoff of the high type can be written  $\phi(\rho, H) = [H - p_1 + \delta(H - L)]\rho$ , while the payoff of the low type can be written as  $\phi(\rho, L) = (L - p_1)\rho$  (the low type receives zero surplus in the second period). Note that  $\partial\phi(\rho, H)/\partial\rho > \partial\phi(\rho, L)/\partial\rho$ , hence  $\phi(\rho, \theta)$  has SCP. Intuitively, the high type in a type 2 equilibrium would be more eager to buy in the first period since he values the object more and cares more about low price in the second period. By MCS Theorem ?? (a), we must have  $\rho_H \geq \rho_L$  in equilibrium. But then by Bayes rule we must have  $\mu_2(1) \geq \mu > L/H$ , and therefore  $p_2(1) = H$  - a contradiction. Therefore, type 2 equilibria do not exist.

**Case 3** In this case we must have  $\mu_2(0) \leq L/H$  and  $\mu_2(1) \geq L/H$ . It can be seen that in this case the optimal equilibrium for the principal is the Full Revelation Equilibrium considered above.

**Case 4** In this case, we must have  $\mu_2(0) \geq L/H$  and  $\mu_2(1) \geq L/H$ . For the former we must have  $\rho_H < 1$ , i.e., some high types not buying in the first period. Note that since the agent's first-period choice does not affect the second-period price, the high type is willing to pay  $H$  and the low type is willing to pay  $L$  in the first period. Thus, for  $\rho_H < 1$  we must have  $p_1 \geq H$ . But then the low type will not buy in the first period, i.e.,  $\rho_L = 0$ .

It will be optimal for the principal to set  $p_1 = H$  and maximize  $\rho_H$  subject to  $\mu_2(0) \geq L/H$ . This gives a "partial revelation" equilibrium. In this equilibrium the principal will maximize the probability of separation of the high type subject to there remaining sufficiently many unrevealed high types so that she can credibly promise not to lower the price. By Bayes' rule,

$$\mu_2(0) = \frac{(1 - \rho_H)\mu}{(1 - \rho_H)\mu + (1 - \mu)}$$

since  $\rho_L = 0$ . The inequality  $\mu_2(0) \geq L/H$  is equivalent to

$$\Leftrightarrow \rho_H \leq \rho^* = \frac{\mu H - L}{\mu(H - L)}.$$

(It is easy to see that  $\rho^* < 1$ .) The principal maximizes expected profit by setting  $\rho_H = \rho^*$  in this candidate equilibrium, which yields profit

$$\pi_P = \mu[\rho_H H + \delta H] = \left[ \frac{\mu H - L}{H - L} + \mu\delta \right] H.$$

**Optimal PBE** When  $\delta < 1$ , Full Revelation is Feasible, and  $\pi_F > \pi_N$ . It remains to compare  $\pi_F$  to  $\pi_P$ , the profits from Partial Revelation:

$$\pi_P \geq \pi_F$$

$$\begin{aligned} \Leftrightarrow \quad & \mu\rho^* H + \mu\delta H \geq \mu H + \delta L \\ \Leftrightarrow \quad & \mu(\delta + \rho^* - 1) \geq \delta \frac{L}{H} \\ \Leftrightarrow \quad & \mu\delta - \frac{(1 - \mu)L}{H - L} \geq \delta \frac{L}{H} \\ \Leftrightarrow \quad & \mu(\delta H - \delta L + L) \geq L \left( 1 + \delta - \delta \frac{L}{H} \right) \\ \Leftrightarrow \quad & \mu \geq \frac{L}{H} \underbrace{\left[ \frac{H + \delta(H - L)}{L + \delta(H - L)} \right]}_{>1} \end{aligned}$$

**Conclusion for ST contracts when  $\delta < 1$ :**

1. If  $\mu < L/H$  then  $P$  implements the full commitment optimum with  $p_1 = p_2 = L$ , and there is no revelation of information.
2. If  $L/H < \mu < \frac{L}{H} \frac{H + \delta(H - L)}{L + \delta(H - L)}$ , then we get the "Full Revelation" equilibrium:

$$\begin{aligned} p_1 &= (-\delta)H + \delta L, \\ p_2(1) &= H \text{ and } p_2(0) = L. \end{aligned}$$

- The  $H$  type buys in both periods.
- The  $L$  type buys only in the second period.

3. If  $\mu > \frac{L}{H} \frac{H + \delta(H - L)}{L + \delta(H - L)}$ , then we get the "Partial Revelation" equilibrium:

$$p_1 = p_2 = H.$$

- The  $H$  type buys in the first period with probability  $\rho^* = \frac{\mu H - L}{\mu(H - L)} < 1$ , and buys for sure in the second period.

- The  $L$  type never buys.

Conversely, when  $\delta > 1$  we know that Full Revelation is infeasible. It remains to compare  $\pi_N$  and  $\pi_P$ :

$$\pi_P - \pi_N = [\mu\rho^*H + \mu\delta H] - [L + \mu\delta H] = \mu\rho^*H - L = \frac{\mu H - L}{H - L}H - L \geq 0$$

$$\Leftrightarrow \mu - \frac{L}{H} \geq \frac{L}{H} \left(1 - \frac{L}{H}\right)$$

$$\Leftrightarrow \mu \geq \frac{L}{H} \left(2 - \frac{L}{H}\right).$$

**Conclusion for ST contracts when  $\delta > 1$ :**

1. If  $\mu < L/H$ , then  $P$  implements the full commitment optimum with  $p_1 = p_2 = L$ , and there is no revelation of information.
2. If  $L/H < \mu < \frac{L}{H} \left(2 - \frac{L}{H}\right)$ , then we get the "No Revelation" equilibrium:

$$p_1 = L, p_2 = H.$$

- The  $H$  type buys in both periods.
- The  $L$  type buys only in the first period.

3. If  $\mu > \frac{L}{H} \left(2 - \frac{L}{H}\right)$ , then we get the "Partial Revelation" equilibrium:

$$p_1 = p_2 = H.$$

- The  $H$  type buys in the first period with probability  $\rho^* = \frac{\mu H - L}{\mu(H - L)} < 1$ , and buys for sure in the second period.
- The  $L$  type never buys.

## Long Term Renegotiable Contracts

Now consider the case where the principal can sign a long-term contract with the agent, but cannot commit not to renegotiate the contract. The game between the principal and the agent can then be described as follows:

- Principal offers a contract  $\langle p_1, p_2(0), p_2(1) \rangle$

- Agent chooses  $x_1$
- Principal offers  $p'_2$
- Agent accepts  $p'_2$  or rejects and sticks to the original contract
- Agent chooses  $x_2$

We will look for the principal's preferred PBE of this game. Note that a renegotiation offer  $p'_2$  following the agent's choice  $x_1$  will be accepted by the agent if and only if  $p'_2 \leq p_2(x_1)$ . Therefore, a long-term renegotiable contract commits the principal against raising the price, but does not commit her against lowering the price.

Analysis is simplified by observing that the principal can without loss offer a contract that is not renegotiated in equilibrium.

**Definition 16.2.1** *A Renegotiation-Proof (RNP) contract is one that is not renegotiated in the continuation equilibrium.*

**Proposition 16.2.1 (Renegotiation-Proofness Principle)** *For any PBE outcome of the game, there exists another PBE that implements the same outcome and in which the principal offers a RNP contract.*

**Proof.** If the principal offer  $\langle p_1, p_2(0), p_2(1) \rangle$  and in equilibrium renegotiates to  $p'_2 \leq p_2(x_1)$  after the agent's choice of  $x_1 \in \{0, 1\}$ , then the contract  $\langle p_1, p'_2(0), p'_2(1) \rangle$  is RNP.

This result is quite general. The general idea is that since the process of renegotiation is a game (mechanism), we can think the "grand mechanism" consisting of the contract and the subsequent renegotiation game. The "grand mechanism" itself is Renegotiation-Proof and implements the same outcome as the original contract followed by renegotiation.

Once we restrict attention to RNP mechanisms, it is clear that the principal's inability to commit not to renegotiate imposes an additional Renegotiation-Proofness constraint on the set of implementable outcomes (in addition to IC and IR constraints). That is, the principal cannot do better, and often does worse, than with full commitment.

**Example 16.2.1** *When  $\mu > L/H$ , with full commitment the principal offers  $p_1 = p_2(0) = p_2(1) = H$ , and the agent chooses  $x_1(H) = x_2(H) = 1$ ,  $x_1(L) = x_2(L) = 0$ . However,*

this outcome is not RNP. Indeed, since  $\mu_2(0) = 0$ , following  $x_1 = 0$  the principal will offer  $p'_2 = L$ , and the agent will accept. Thus, the RNP constraint reduces the principal's payoff.

The next observation is that the principal is better off with  $L$  renegotiable contracts than with ST contracts. Indeed, in the former case one possibility for the principal is to offer an ex ante contract specifying the default outcome of the second period. In our example, this corresponds to  $p_2(0) = p_2(1) = \infty$ . This will implement the ST contracting outcome. Thus we have

**Proposition 16.2.2** *If an outcome is implementable with ST contracting then it is also implementable with a LT Renegotiable contract (and therefore, by the RNP Principle, with a LT RNP contract).*

In our example, the intuition is that with LT renegotiable contracts the principal cannot commit against reducing the 2nd-period price, while with ST contracting she cannot commit against any modification of the 2nd-period price.

When do long-term renegotiable contracts offer an advantage relative to short-term contracts? When the ability to commit to a low price is useful. For example, consider the possibility of full revelation in the first period when  $\delta > 1$ . With short-term contracts, full revelation was impossible: because of the ratchet effect, the high type would only buy in the first period at a price  $p_1 < L$ , but at this price the low type would also buy - "take the money and run".

It turns out, however, that the principal can achieve full revelation with a LT RNP contract  $< p_1, p_2(0), p_2(1) >$ . In particular, consider the following contract:

$$\begin{aligned} p_1 &= H + \delta L, \\ p_2(1) &= 0, \\ p_2(0) &= L. \end{aligned}$$

This contract will have a full-revelation equilibrium. Indeed,  $(IC_{HL})$  is satisfied:  $H$ 's surplus is  $\delta(H - L)$  regardless of whether he buys or not.  $(IC_{LH})$  is also satisfied:  $L$ 's surplus would be negative if he bought. Intuitively, unlike in the ST contracting case, both incentive constraints can now be satisfied by committing against ratcheting - setting

$p_2(1) = 0$ . This allows to make the "take the money and run strategy" costly for the low type - he now has to pay for the 2nd-period consumption as well. He will not be willing to pay price  $H + \delta L$  for both periods' consumptions, while the high type will pay it. The contract is also RNP, since  $p_2(x_1) \leq L$  for all  $x_1$ .

Hart and Tirole (1988) reinterpret the model as describing contracting about the use of durable good. A short-term contract in this setting is a rental contract. The long-term contract with  $p_2(1) = 0$  can be interpreted as a sales contract. Thus, by using sales contracts instead of rental contracts, the seller eliminates the ratchet effect, make full revelation of information possible (but not always optimal). The renegotiation effect remains.

More generally, Hart and Tirole consider a multi-period model. They show that long-term renegotiable rental contracts give rise to the same outcome as sales contracts. The resulting price follows a Coasian Dynamics: it decreases over time and eventually drops to  $L$ . When the number of periods is large and  $\delta$  is close to 1, the drop is almost immediate in real-time (cf. Gul-Sonnenschein-Wilson 1988, Coase 1972).

## Reference

Bolton, P. and M. Dewatripont (2005), *Contract Theory*, MIT Press.

Mas-Colell, A., M. D. Whinston, and J. Green, *Microeconomic*, Oxford University Press, 1995, Chapter 14.

Segal, I. (2007), *Lecture Notes in Contract Theory*, Stanford University.

Wolfstetter, E., *Topics in Microeconomics - Industrial Organization, Auctions, and Incentives*, Cambridge Press, 1999, Chapters 8-10.